

1498

a.h. boon van ostade

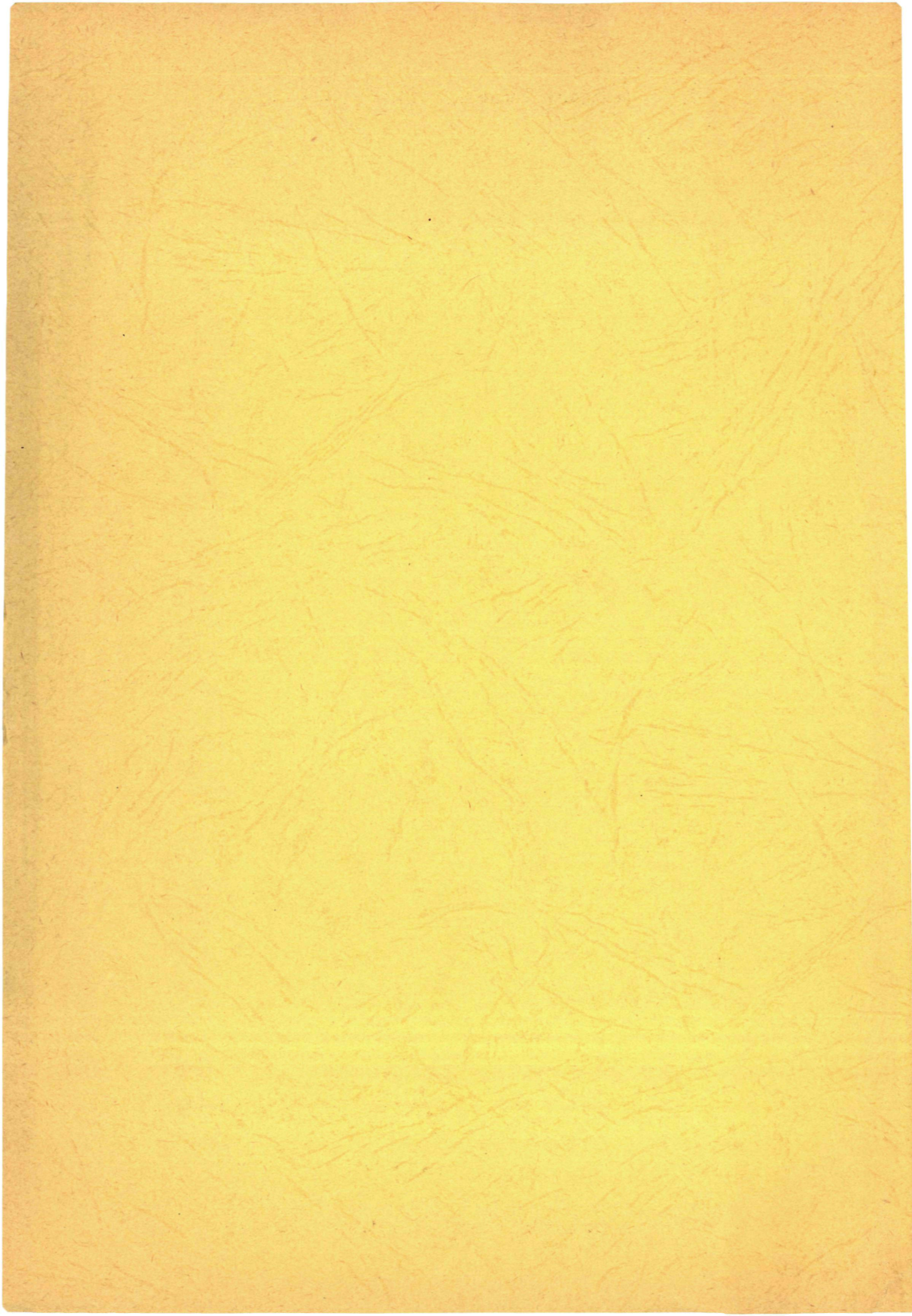
d
e
i
t
e
r
a
t
i
e
e
v

c
l
u
s
t
e
r
a
n
a
y
l
s
e

een klassificatiemethode

voor

psychologische data



DE ITERATIEVE CLUSTERANALYSE

Een klassificatiemethode voor psychologische data

Promotor :

Prof. Dr. G. Brenninkmeijer

DE ITERATIEVE CLUSTERANALYSE

Een klassificatiemethode voor psychologische data

Proefschrift

ter verkrijging van de graad van
doctor in de sociale wetenschappen
aan de katholieke universiteit te Nijmegen
op gezag van de rector magnificus,
Dr. G. Brenninkmeijer, hoogleraar in
de faculteit der sociale wetenschappen
volgens besluit van de senaat
in het openbaar te verdedigen
op 16 december 1969, des namiddags te 4 uur

door

Adriaan Herman Boon van Ostade

geboren te Bandoeng

Dank zij de hulp en steun van zeer velen is deze studie tot stand gekomen. Het is mij onmogelijk hier allen persoonlijk te bedanken.

Voor drie personen moet echter een uitzondering worden gemaakt en wel voor Prof. Dr. A.D. de Groot, Prof. Dr. S.D. Fokkema en Dr. G.B.M.L. Koene. Alle drie hebben zij de studie op belangrijke momenten verder voortgeholpen.

| | pagina |
|--------------------------------------------------------|--------|
| 1. Inleiding | 7 |
| 2. Klassificatie in de psychologie | |
| 2.1. Algemeen logisch uitgangspunt | 9 |
| 2.2. De elementen | 12 |
| 2.3. De operator | 13 |
| 2.4. Het doel van de klassificatie | 17 |
| 2.5. De orde van de klassificatie | 22 |
| 3. Constructie van doelmatige klassificaties | |
| 3.1. De waarde van de operator | 25 |
| 3.2. De bedoelde klassificatie | 35 |
| 3.3. De klassificatiemaat | 46 |
| 3.4. Constructievoorschriften | 53 |
| 3.5. Controle en exploratie | 61 |
| 3.6. De waarde van de geconstrueerde klassificatie | 68 |
| 4. Uitvoering van de analyse | |
| 4.1. De steekproef | 73 |
| 4.2. De operatoren | 79 |
| 4.3. Omzetting van de data in nul en één coderingen | 83 |
| 4.4. De overeenkomstindex | 90 |
| 4.5. Vorming van de clusters | 94 |
| 4.6. Controle van de clusters | 104 |
| 4.7. Rekenprogramma | 113 |
| 5. Toepassingen van de analyse | |
| 5.1. Inleiding | 125 |
| 5.2. Soorten van toepassing | 126 |
| 5.2.1. Betekenisonderzoek | 129 |
| 5.2.2. Psychische functies | 139 |
| 5.2.3. Interindividueelonderzoek | 143 |
| 5.2.4. Type-onderzoek | 144 |

| | |
|-------------------------------------------------------------|-----|
| 3. Vormen van toepassing | 150 |
| 5.3.1. Conditie-onderzoek | 152 |
| 5.3.2. Hiërarchisch-onderzoek | 159 |
| 5.3.3. Predictie-onderzoek | 162 |
| 6. Vergelijking met de hoofdassenanalyse | |
| 6.1. Inleiding | 167 |
| 6.2. Kritische punten hoofdassenanalyse | 175 |
| 6.2.1. Het aantal factoren | 181 |
| 6.2.2. Aselecte data | 184 |
| 6.3. Kritische punten iteratieve clusteranalyse | 188 |
| 6.3.1. De ladingen | 193 |
| 6.3.2. Subclusters | 197 |
| 7. Samenvatting | 201 |
| 8. Summary | 205 |
| 9. Literatuur | 209 |
| Bijlagen | |
| I. Statistische bewerkingen | |
| I.1. De betrouwbaarheid van de point-biserial r | |
| I.2. Toetsing van een reeks afhankelijke r -waarden | |
| I.3. Toetsing van een reeks afhankelijke p -waarden | |
| I.4. Toetsing van de K.R. (21) en K.R. (20) | |
| I.5. Omzetting van de point-biserial r in de biserial r | |
| II. De data van de behandelde onderzoeken | |
| II.1. Data simulatie-onderzoeken (zie paragraaf 6.1.) | |
| II.2. Data zusters (zie paragraaf 4.1.) | |
| II.3. Data beroepen (zie paragraaf 5.2.1.) | |
| II.4. Data opstellen (zie paragraaf 5.2.4.) | |
| II.5. Data A.M. onderzoek (zie paragraaf 4.5.) | |
| III. Populatie factorstructuur simulatie-onderzoeken | |

1. Inleiding

Wanneer in 1969 een studie in de psychologie wordt geschreven over een klassificatiemethode, is het bijna niet mogelijk dit zonder technische verantwoording te doen. Het accent op deze technische aspecten zou de indruk kunnen wekken dat het om een theoretische studie gaat. Dit is echter niet juist. De publicatie is vooral bedoeld voor de praktijk van het toegepast onderzoek. Daarom zijn theoretische opvattingen en bedenkingen in deze studie regelmatig getoetst op hun praktische relevantie. Dit criterium wordt zonder verdere uitleg gehanteerd. Het is daarom goed hier te omschrijven wat daaronder wordt verstaan.

Praktisch niet-relevant worden al die bedenkingen genoemd, die geen of nauwelijks enige aanleiding kunnen geven tot foutieve interpretaties, wanneer met deze bedenkingen geen rekening wordt gehouden. Zo is voor een correlatie coëfficiënt (de in deze studie meest gebruikte statistische maat) een afwijking in de derde decimaal praktisch meestal niet-relevant. Dergelijke verschillen zijn namelijk bij correlaties van 0,75 of kleiner pas bij steekproeven van tienduizend of meer elementen statistisch generaliseerbaar.

De studie zelf bestaat uit vijf hoofdstukken. Het eerst geeft de uitgangspunten en definities onder de algemene titel 'klassificatie in de psychologie'. De belangrijkste paragrafen van dit hoofdstuk zijn die welke gaan over 'de operator' en 'het doel van de klassificatie'. Het tweede hoofdstuk is een nadere theoretische uitwerking voor de constructie van doelmatige klassificaties. De daarop volgende twee hoofdstukken vormen de handleiding van de methode met voorbeelden van toepassing. In paragraaf 4.7. wordt de rekenmethode van de analyse gegeven. Het laatste hoofdstuk

is een verslag van een uitgevoerde vergelijking van de iteratieve clusteranalyse met de hoofdassenanalyse bij acht simulatie-onderzoeken.

De psycholoog wordt dagelijks in zijn werk geconfronteerd met een onoverzichtelijke hoeveelheid van verschijnselen. Nu tot op een bepaalde hoogte de observatie-problematiek in de psychologie is opgelost, wordt het noodzakelijk doelmatige ordeningen aan te brengen in deze verschijnselen. Als de iteratieve clusteranalyse een reële bijdrage blijkt te zijn voor de constructie van deze klassificaties, heeft deze studie beantwoord aan zijn doelstelling.

2. Klassificatie in de psychologie

2.1. Algemeen logisch uitgangspunt

De logica heeft zich al sinds Aristoteles bezig gehouden met de theorie der klassen (Mil, z.j., p.76 e.v.). In deze theorie zijn de begrippen 'genus', 'species', 'differentia', 'proprium' en 'accidens' ontwikkeld.

Met genus en species wordt uitgedrukt dat iedere klasse een genus of soortbegrip is voor een aantal subklassen en tegelijkertijd een subklasse of species van een hogere klasse of genus superior. De kenmerken waarmee een klasse kan worden onderscheiden van andere klassen worden differentia genoemd. Subklassen, hogere klassen en nevenschikkende klassen verschillen verder op grond van eigenschappen of propria. De eigenschappen moeten essentieel zijn voor de klasse, niet toevallig, geen accidens zijn. Een belangrijk gedeelte van de oude logica gaat over het verschil tussen essens en accidens. Met Mill zal onder dit verschil "which has given occasion to so much abstruse speculation, and to which so mysterious a character was formerly, and by many writers is still, attached" niets meer worden verstaan "than the difference between those attributes of the class which are, and those which are not, involved in the signification of the class-name" (p.78). Wel maakt Mill een onderscheid tussen meer of minder toevallige of min of meer essentiële klassificaties en wel afhankelijk van het aantal kenmerken waarin de klasse verschilt van andere klassen. Hoe groter dit aantal kenmerken, des te essentiëler is de klassificatie.

De moderne logica heeft met de ontwikkeling van de logistiek bovengenoemde onderscheidingen verlaten en is gekomen met een begrip-penkader, dat een nadere uitwerking is van bovengenoemde onder-

scheidingen. Het is de Duitse wiskundige Cantor (zie Tarski, 1964, p.74) die met de leer der verzamelingen hiertoe veel heeft bijgedragen. Een klasse wordt opgevat als een verzameling van elementen. Dit is een klasse van de eerste orde. Klassen van de tweede orde bestaan niet uit elementen maar uit klassen van de eerste orde. Op die wijze kunnen klassen van derde, vierde orde worden onderscheiden. Zo kunnen de passagiers van een vliegtuig worden opgevat als elementen van een klasse van de eerste orde. Klassen van vliegtuigen zijn dan klassen van de tweede orde. Klassen van vervoermiddelen van de derde orde enz.

Van fundamentele betekenis in de theorie der klassen zijn zinswendingen als *individu b behoort tot de klasse K, de klasse K bevat het ding d als element (of lid)*. Deze uitdrukkingen worden aangegeven met de volgende symbolen

$$b \in K \quad \text{of} \quad d \in K$$

De elementen (b,d) worden opgevat als vrije veranderlijken of zaken, die niet constant zijn. Constanten zijn termen, symbolen, die een vaste betekenis hebben welke in de loop van de beschouwing niet verandert. De etta ϵ is een voorbeeld van een constante, echter ook het getal nul en een bepaald observatie- of coderingsvoorschrift. Een veranderlijke heeft op zichzelf genomen geen vaste betekenis. Een klassificatie bindt nu een veranderlijke, geeft hem betekenis. Een klassificatie is daarom een operator, ook wel een quantor genoemd. Zo kan een element worden gebonden door het een mens te noemen. Verdere klassificaties als bijvoorbeeld: een bezoeker, een kind, Jan Teulings, geven steeds meer specificaties en binden de veranderlijke steeds meer. Met de naamgeving is een vrije veranderlijke alleen nog

mogelijk door personen met dezelfde naam of door veranderingen van dezelfde persoon in de tijd 'Jan is zichzelf niet vandaag'.

De elementen van een klasse moeten aan een bepaalde eigenschap (of eigenschappen) voldoen opdat zij tot de klasse kunnen worden gerekend. Om die reden kan de theorie der klassen tot een bepaalde hoogte gelijk worden gesteld aan de theorie der eigenschappen (Tarski, 1964, p.77). Een klasse wordt bepaald door de uitdrukking

de klasse van alle elementen b die ...

Zo kan de klasse van kinderen worden aangegeven als de verzameling van alle mensen die niet ouder zijn dan vijftien jaar.

Klassen kunnen verder elkaar geheel omvatten, overlappen of geheel uitsluiten en wel afhankelijk van het feit of de elementen dezelfde zijn, gedeeltelijk dezelfde zijn of geheel verschillen. De notaties hiervoor zijn:

$K_1 = K_2$ (alle elementen van K_1 zijn de elementen K_2 en omgekeerd)

$K_1 \subset K_2$ (K_1 is een deelklasse van K_2),

$K_1 \supset K_2$ (K_2 is een deelklasse van K_1) en

$K_1 \neq K_2$ (K_1 en K_2 hebben geen enkel element gemeenschappelijk)

Onder een universele klasse wordt verder verstaan die klasse, waar alle elementen invallen, waarover in de context van een bepaalde theorie wordt gehandeld. Aangegeven met $b \in V$. Iedere klasse is nu een deelklasse van de universele klasse $K \subset V$.

2.2. De elementen

Het genus van de psychologie zal in deze studie worden opgevat als het levende, niet plantaardige organisme, te onderscheiden in de species dieren en mensen.

Elementen of vrije veranderlijken in de psychologie zijn in eerste instantie; organismen, ook wel individuen of personen genoemd. Klassificaties van individuen zullen classificaties van de eerste orde worden genoemd. Het is ook mogelijk hetzelfde individu in de tijd als variabel op te vatten. Dit zijn dan intra-individuele classificaties. Bij classificaties van verschillende personen zijn het inter-individuele. Klassificaties van groepen van individuen, zoals gebruikelijk in de sociale psychologie zijn dan classificaties van de tweede orde.

Naast individuen komen als vrije veranderlijken in de psychologie 'stimuli' voor. Klanken, kleuren, vragen, dingen, be-roepen kunnen 'stimuli' vormen. Ook gedrag van personen kan als stimulus worden opgevat. Klassen van stimuli zijn van de eerste orde. Een stimulus of informatie-eenheid komt echter als apart element meestal niet als zodanig voor. Het gaat vaak om prikkelpatronen of informatie-structuren; om verzamelingen van stimuli. Klassen van stimulusconfiguraties zijn klassen van de tweede orde.

Met Duyker (1959, p.200) kan het primaire object der psychologie worden omschreven als: het menselijk gedrag. Hieruit volgt dat een classificatie in de psychologie altijd zou plaats vinden op grond van gedrag. Individuen en stimuli zouden geklassificeerd worden op grond van hun gedrag of resultaten van gedrag. Om die reden worden gedragingen niet als mogelijke elementen opgevat.

2.3. De operator

Om klassificaties uit te voeren kunnen schema's worden gebruikt. Een dergelijk schema kan bestaan uit een observatievoorschrift waarin wordt aangegeven hoe de observaties moeten worden geregistreerd en een coderingsvoorschrift dat aangeeft bij welke uitkomsten van de observatie het element in de klasse valt.

Het observatievoorschrift kan aanwijzingen omvatten over de eventuele te gebruiken apparatuur, vragen, opdrachten enz. Zo wordt de lengte van rekruten (de elementen) bij de keuring op een bepaalde wijze vastgelegd (de observatie) om te bepalen wie groter is dan bijv. 1.90 en kleiner dan 1.65 (codering). Deze rekruten worden afgekeurd (doel van de klassificatie).

Coombs (1964, p.4 en 5) spreekt in dit verband over twee fasen. In de eerste fase worden 'recorded observations' verkregen, in de tweede 'data'. Fase 2 formuleert hij als volgt: "these recorded observations, however, are not yet data in the sense of this theory of data; an interpretive step on the part of the scientist, called phase 2 is required to convert the recorded observations into data. Phase 2 involves a classification of observations in that sense that individuals and stimuli are identified and labeled".

Bovenstaande kan als volgt worden geformaliseerd, waarmee de operator wordt gedefinieerd.

$F(b) = 0$ (observatievoorschrift) en

als $0 = W$ dan en dan alleen $b \in K$ (coderingsvoorschrift)

Hierbij is b het element, 0 de observatie, W de verwachte observatie en K de klasse. Het is daarbij niet nodig dat 0 een bepaald getal is

0 kan een kwaliteit of discrete variabele zijn. Zo kan voor de klassificatie van een kleur het observatie- en coderingsvoorschrift bestaan uit de opdracht aan een aantal beoordelaars om een kleur (de observatie: 0) van een bepaald object (het element: b) te vergelijken met een vaststaande kleur (W). Als de kleur, volgens de beoordelaars overeen komt met de gegeven kleur W, dan valt het object in de klasse K.

Het observatievoorschrift is geformaliseerd als een functie van b. Hiermee wordt tot uitdrukking gebracht, dat met element b een aantal handelingen kan worden verricht om 0 te krijgen.

Het is noodzakelijk erop te wijzen dat de observatie- en coderingsvoorschriften van een operator in deze beschouwing als constant worden opgevat. Wanneer de voorschriften variabel worden toegepast zijn de klassificaties van verschillende elementen niet meer onderling vergelijkbaar. De reden, dat dan element b in de klasse valt en element d niet, kan dan zowel aan een verschil tussen de elementen als aan een verschil tussen de voorschriften worden geweten. Uiteraard kunnen voorschriften nooit strikt constant worden toegepast. Buitenexperimentele variabelen hebben altijd een storende werking. Operatoren zijn beter naarmate zij voor dergelijke fouten minder gevoelig zijn.

Daarnaast dienen operatoren ten opzichte van elkaar experimenteel onafhankelijk te zijn. Met de resultaten van operator i mag niet zodanig experimenteel worden gemanipuleerd dat hierdoor de resultaten van operator j worden beïnvloed. De term 'experimenteel' is hierbij belangrijk. De resultaten van de operatoren kunnen onderling

wel overeenkomen. Dit kan ook theoretisch worden verwacht, bijvoorbeeld omdat gesteld wordt dat de uitslagen door gelijke condities worden bepaald. Het onderzoek of experiment mag echter niet zodanig worden uitgevoerd, dat deze overeenkomst noodzakelijkerwijs moet optreden. Uit de data moet blijken, dat de veronderstelling empirisch beantwoord aan de verwachting. Het moet mogelijk zijn geen overeenstemming te vinden. De falsifieerbaarheid van de verwachting mag niet door een bepaalde experimentele opzet worden uitgesloten. Om die reden zijn bijvoorbeeld items van een zuivere 'speed'-test niet als onafhankelijke operatoren te beschouwen. De 'speed'-test is één operator. De term 'klasse' wordt verder vaak gebruikt om ordinale of getalsmatige uitslagen te benoemen; bijvoorbeeld 'deze grafische schaal heeft zeven klassen' of 'de score is verdeeld in tien klassen'. Dergelijke klassificaties zijn ook niet experimenteel onafhankelijk. Als element b in één van dergelijke klassen valt, komt dit element noodzakelijkerwijs niet in de andere klassen. In verschil met het operationele klassebegrip, zoals hier wordt gehanteerd, zullen dergelijke klassen 'eenheden' worden genoemd. De verschillende antwoordcategorieën van meer-keuze items worden ook vaak als klassen aangeduid. Ook deze klassen zijn experimenteel afhankelijk. In deze studie zullen dergelijke kwalitatieve onderscheidingen worden aangegeven met de term 'categorieën'.

Operatoren worden vaak 'variabelen' genoemd. Uit het voorafgaande is het duidelijk, dat hier onder moet worden verstaan: operaties die variatie (onderscheidingen) aanbrengen tussen elementen. De operatoren zelf zijn niet variabel.

Uit bovenstaande volgt dat in de meeste studies een beoordelaar, observator of proefleider als een onderdeel van de operator moet worden opgevat. Een afhankelijkheid van de resultaten van de operator van de

persoonlijkheid of identiteit van de beoordelaar maakt de operator subjectief. Sommige projectietests worden op deze wijze vaak toegepast. Hoewel op zichzelf tegen een dergelijk gebruik voor het nemen van praktische beslissingen weinig bezwaar kan worden gemaakt, zijn zulke operatoren moeilijk te controleren, te verbeteren of over te dragen. De uitspraken gebaseerd op genoemde procedures hebben daarom wetenschappelijk gezien een beperkte betekenis. Het hangt van het doel van het onderzoek af of gesproken kan worden over beoordelaars dan wel proefpersonen. Bij bestudering van de oordelen, de functie van het oordelen zijn de beoordelaars 'proefpersonen', wanneer de oordelen echter gebruikt worden om de mensen of stimuli, die beoordeeld zijn, te bestuderen vormen de beoordelaars een onderdeel van de operator.

2.4. Doel van de klassificatie

Het doel van klassificatie van elementen is gelegen in de mogelijkheid om tot deducties te komen over de elementen, die in de klasse vallen. Dergelijke deducties hebben de vorm van een implicatie

als b in klasse K valt dan kenmerk k ($b \in K \rightarrow k$)

Het eerste lid van de uitspraak wordt het onderstelde genoemd, het tweede het gestelde.

Dergelijke deducties zijn echter alleen doelmatig wanneer ze niet opgaan voor de elementen, die niet in de klasse vallen;

als b niet in klasse K valt dan niet kenmerk k

In formule luidt de inversie

$b \notin K \rightarrow \sim k$

volgens de wet der contrapositie (Tarski, 1964, p.47 t/m 50) is de inversie gelijkwaardig aan de inductie

$k \rightarrow b \in K$

als k dan valt b in K

Wanneer de inductie en deductie tegelijk waar moeten zijn, zijn gestelde en onderstelde gelijkwaardig en is de uitdrukking 'dan en dan alleen' toepasbaar. Deze uitdrukking wordt ook gebruikt voor definities. Een kenmerk van een klasse is dus een definitie van een klasse dan wel een eigenschap van de elementen die in de klasse vallen en niet van de elementen die niet in de klasse vallen. Op dit punt zijn de theorieën der klassen, der eigenschappen en der definities overeenkomstig.

In formule

$$(1) \quad b \in K \leftrightarrow k \quad \text{of} \quad K \leftrightarrow k$$

Een klassificatie is dus doelmatig wanneer de implicaties van de klasse zowel deducties als inducties zijn; definities zijn.

Bovenstaande kan met een voorbeeld nader worden geïllustreerd. Kinderen kunnen geklassificeerd worden: 'in bang om boodschappen te doen' en 'niet bang'. Gededuceerd kan nu worden, dat kinderen die bang zijn om boodschappen te doen zich inefficiënter zullen gedragen bij het doen van boodschappen dan kinderen die niet bang zijn. Wanneer deze deductie ook een inductie is voor de klassificatie, zou ook gezegd kunnen worden, dat kinderen die zich inefficiënt gedragen bij het doen van boodschappen hiervoor bang zijn.

Dergelijke tweezijdige implicaties worden in de werkelijkheid maar in een bepaalde mate gevonden. Een klassificatie is echter des te doelmatiger des te meer deducties tegelijkertijd ook inducties zijn.

Met John Stuart Mill (z.j.) is in de aanvang van deze studie reeds gesteld dat een klassificatie des te essentiëler is des te meer kenmerken opgaan voor de elementen van de klasse. Een klassificatie is dus essentiëler des te meer deducties in bovengenoemde zin mogelijk zijn over de klasse-elementen. Er moet dus gestreefd worden naar een set van kenmerken waar formule 1 voor opgaat.

Een klassificatie is niet uitvoerbaar als er niet minstens één operator is om de klassificatie uit te voeren. De operator werd in de vorige paragraaf als volgt gedefinieerd:

$$f(b) = 0 \wedge 0 = W \leftrightarrow b \in K$$

waarbij b het element is,

0 de observatie,

W de verwachte observatie, en

K de klasse

Een eenvoudiger notatie geeft

$$(2) \quad \underline{0} \leftrightarrow K$$

waarbij $\underline{0}$ de operator en K de klasse is. Een doelmatige en essentiële klassificatie is dus in formule

$$(3) \quad \underline{0} \leftrightarrow K \leftrightarrow k_1 \wedge k_2 \dots \wedge k_i \dots \wedge k_n$$

Een empirische controle op de juistheid van deze klassificatie kan nu plaats vinden door de kenmerken om te zetten in operatoren en na te gaan in hoeverre de klassificaties overlappen. De set van operatoren $\underline{0}(k_1), \underline{0}(k_2), \dots, \underline{0}(k_i), \dots, \underline{0}(k_n)$ is dan representatief voor K .

Empirisch wordt nagegaan in hoeverre de klassen $K_1, K_2, \dots, K_i, \dots, K_n$ elementen gemeenschappelijk hebben. Om twee redenen kan dit onvoldoende zijn, of omdat formule 3 niet juist is of omdat de specificatie van de kenmerken niet goed heeft plaats gevonden. Methodisch is waarschijnlijk de meest aanvaardbare procedure om dit laatste aan te nemen en pas op de duur, wanneer bij herhaling de specificatie mislukt formule 3 als onjuist te beschouwen. In de volgende hoofdstukken wordt dit verder uitgewerkt.

Wanneer *alle* elementen (stimuli of proefpersonen) in de klasse vallen is de klassificatie ondoelmatig; de klassificatie een universele klasse zou zijn, als $K = V$. In de psychologie is dat de klasse 'mens' of 'levend organisme' en de klasse 'stimuli'. De klassen moeten subklassen

zijn van deze universele klassen. Om die reden is de uitspraak 'alle mensen zijn sterfelijk' alleen doelmatig in die opvattingen, waar ook de mogelijkheid van onsterfelijke wezens als mogelijk waar wordt beschouwd. De nul klasse is ook niet doelmatig. Alle elementen, die niet in de klasse vallen, vormen namelijk weer een universele klasse.

Optimaal is verder de klassificatie van een operator die 50% van de elementen in de klasse laat vallen. In dat geval geeft de operator de meeste informatie over de verzameling van elementen. Dit is af te lezen uit Shannon's formule voor informatie;

$$H = \sum_{i=1}^n p_i \cdot \log_2 1/p_i$$

Hierbij is H de hoeveelheid informatiegehalte,

p_i de kans op boodschap i, en

n het totaal aantal onafhankelijke boodschappen.

Bij een operator, zoals gedefinieerd in paragraaf 2.3., is $n=2$: 'valt in de klasse' (1), 'valt niet in de klasse' (2). H krijgt nu zijn grootste waarde als $p_1=p_2=0,50$ en is nul als $p_1=1,00$ of $0,00$ (de universele- resp. nul klasse).

Samenvattend kan worden gezegd dat een klassificatie des te doelmatiger (essentiëler en informatief) is als

- a. het aantal eigenschappen of kenmerken van de klasse zo groot mogelijk is,
- b. deze eigenschappen zowel inductief als deductief kunnen worden toegepast op de elementen, en
- c. er minstens één operator is die zoveel mogelijk vijftig procent van de elementen in de klasse laat vallen,

terwijl een klassificatie beslist ondoelmatig is als

a. de klassificatie geen operator heeft, en

b. de klassificatie alle of geen elementen bevat.

2.5. De orde van de klassificatie

In de psychologie is de orde van de klassificatie gegeven met het hiërarchische gezichtspunt, het eerst geformuleerd door Burt in 1941 en 1949. Bij deze theorie worden vier niveau's onderscheiden, in overeenstemming met de vier mogelijke typen van factoren bij factoranalyse: algemene, groeps-, specifieke- en foutenfactoren. Guilford (1959, 1961) en Eysenck (1951) zien in de organisatie van het gedrag vier overeenkomstige niveau's: typen, trekken, gewoonte- en specifieke antwoorden.

Coan (1964) wijst erop, dat één van de problemen gevormd wordt door de vraag hoe empirisch gevonden factoren (of klassen) geïnterpreteerd moeten worden in bovengenoemde niveau's. Hij acht het daarom nuttig een onderscheid te maken tussen generaliseerbaarheid naar inhoud en referentie. Bij de eerste gaat het om de empirisch gevonden klassen. Bij de tweede om de referentie vanuit het gedragssysteem: de psychologische theorie.

In aansluiting met de ontwikkelde classificatietheorie kan worden gesteld, dat zowel bij de operatoren als bij de kenmerken een hiërarchie is te onderscheiden. De eerste geeft een generaliseerbaarheid naar inhoud, de tweede naar referentie. Het is niet noodzakelijk dat beide hiërarchieën direct parallel lopen. Het is bijvoorbeeld mogelijk een typenindeling te verkrijgen met één operator en een gewoonte-antwoord als 'zich-goed-voordoen' met een groot aantal vragen of operatoren, die alle laden in één algemene factor. Het aantal operatoren is niet relevant. De typenindeling is echter alleen juist als de kenmerken van het type omgezet in operatoren samen met de gestelde operator tot een bepaalde hoogte één klasse vormen: de elementen in voldoende mate overeenkomstig klassificeren. Wetenschappelijk gezien dient dus in voldoende mate aangetoond te zijn, dat de twee hiërarchieën isomorf zijn.

Johnson (1967) heeft in navolging van Ward (1963) een hierarchisch clusterschema uitgewerkt. Het blijkt daarbij dat de sterkte of waarde van de clustering kan worden geoperationaliseerd met een overeenkomstindex. Zijn schema gaat uit van elementen of objecten. Wij zullen spreken over operatoren, waarvan de klassificaties van de elementen min of meer overlappen. Het verschil is niet principieel. In tabel 2.5.1. staat een hierarchisch clusterschema.

Tabel 2.5.1

Een hierarchisch clusterschema
(naar Johnson, 1967, en Ward, 1963)

| sterkte | overeenkomst | operatoren | | | | | | | | | | | |
|---------|--------------|------------|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | | 3 | | 5 | | 6 | | 4 | | 2 | |
| 0,00 | 1,00 | . | | . | | . | | . | | . | | . | |
| 0,04 | 0,74 | . | | x | x | x | | . | | . | | . | |
| 0,07 | 0,40 | . | | x | x | x | x | x | | . | | . | |
| 0,23 | 0,32 | x | x | x | x | x | x | x | | x | x | x | |
| 0,31 | 0,20 | x | x | x | x | x | x | x | x | x | x | x | x |

Bij het eerste niveau van de hiërarchie is iedere operator een aparte klasse van elementen. De sterkte van de clustering is 0,00. De overeenkomst is volledig (1,00). De klassificatie van operator 1,3,5,6,4 en 2 is identiek met zichzelf; $K_1 = K_1$, $K_2 = K_2$, $K_3 = K_3$, $K_4 = K_4$, $K_5 = K_5$, $K_6 = K_6$. Bij een overeenkomst van 0,74 blijken K_3 en K_5 voldoende te overlappen om samen een cluster te vormen. De clusteringsterkte is 0,04. Bij een overeenkomst van 0,40 kunnen de klassificaties van operatoren 3,5 en 6 als overeenkomstig worden beschouwd. Bij 0,32 de operatoren 1,3,5 en 6 en apart 4 en 2. Bij 0,20 vormen alle operatoren één cluster. De waarde van de clustering is 0,31.

De getallen in het afgebeelde schema dienen alleen om de gedachten te bepalen. Naderhand zal aan de overeenkomstindex op een enigszins andere wijze dan door Johnson een exacte inhoud worden gegeven.

In deze paragraaf is de orde van de klassificatie uitgewerkt voor de operatoren en de kenmerken. Het is ook mogelijk te spreken over de orde van de elementen, zoals onder andere in paragraaf 2.2. plaats vindt. Het verschil is, zoals al gesteld niet principieel. Elementen zijn overeenkomstig naarmate ze meer kenmerken gemeenschappelijk hebben, operatoren naarmate ze dezelfde elementen klassificeren. Op deze kwestie zal in paragraaf 5.3.2. verder worden ingegaan.

3. Constructie van doelmatige klassificaties

3.1. De waarde van de operator

De waarde van een operator heeft verschillende aspecten. Het is mogelijk operatoren te vergelijken voor wat betreft eventuele ethische problemen verbonden aan de uitvoering of voor wat betreft de kosten. Ook kan de psychometrische betrouwbaarheid en de validiteit in de waardering worden betrokken. Vergeleken met de validiteitsvraag zijn echter kwesties van betrouwbaarheid en interne efficiëntie van secundaire betekenis (de Groot, 1961, p.264).

Om die reden zal verder alleen maar over de validiteit van de operator worden gesproken; de vraag in hoeverre de operator een aanvaardbare, adequate methode is voor bepaling van de klasse zoals bedoeld. Eenzelfde operator kan verschillende validiteiten hebben en wel afhankelijk van de klasse, waarvoor hij in een bepaalde onderzoek-context wordt gesteld (de Groot, 1961, p.89 en p.261). De validiteit wordt dus gedefinieerd door de bedoelde klassificatie. In welke mate gaat de relatie van formule 3 voor operator en bedoelde klassificatie op? In hoeverre is

$$\underline{O} \leftrightarrow K \leftrightarrow k$$

waarbij \underline{O} de operator is,

K de bedoelde klasse, en

k het kenmerk waarvan \underline{O} de specificatie is.

De inhoudsvaliditeit van de operator is gewaarborgd als specificatie van de klasse. Empirisch moet echter worden gecontroleerd in hoeverre deze specificatie ook beantwoordt aan de gestelde verwachtingen.

Bij de (begrips)validiteit van de operatoren gaat het nu om de onderdelen $\underline{0}$ en $K \leftrightarrow k$ en hun onderlinge functie 'dan en dan alleen'. De onderdelen kunnen waar (wa) of onwaar (on) zijn. Op grond daarvan zijn er vier mogelijkheden bij klassificatie van een element b door de operator en wel

- a. element b heeft het kenmerk (K is waar) en valt volgens de operator in de klasse ($\underline{0}$ is waar),
- b. element b heeft het kenmerk niet (K is onwaar) en valt volgens de operator in de klasse ($\underline{0}$ is waar),
- c. element b heeft het kenmerk (K is waar) en valt volgens de operator niet in de klasse ($\underline{0}$ is onwaar), en
- d. element b heeft het kenmerk niet (K is onwaar) en valt volgens de operator niet in de klasse ($\underline{0}$ is onwaar).

Van deze vier mogelijkheden is de gestelde functie bij b en c onwaar. Dit kan worden aangetoond met behulp van de calculus der volzinnen (Tarski, 1964, p.41 t/m 50). In tabel 3.1.1. staat de waarheidstafel van deze functie voor deze onderdelen.

Tabel 3.1.1.

Waarheidstafel

| | $\underline{0}$ | $K \leftrightarrow k$ | $\underline{0} \leftrightarrow K \leftrightarrow k$ |
|---|-----------------|-----------------------|-----------------------------------------------------|
| a | wa | wa | wa |
| b | wa | on | on |
| c | on | wa | on |
| d | on | on | wa |

Het volgende voorbeeld van mogelijkheid b kan een en ander verduidelijken. Deze kleuter (b) heeft op de vraag hoeveel twee plus twee is ($f(b)$) het goede antwoord vier ($0=W$) gegeven, toch kan deze kleuter niet rekenen ($\neg k$).

van een bruikbare operator. (Dergelijke operatoren kunnen worden gespiegeld. Een dergelijke operator zit er consequent naast. Het is de spiegel die de waarheid omgekeerd weerkaatst; Parkinson, 1963, p.113 e.v.). Een operator, waar beide kansen nagenoeg gelijk zijn, is onbruikbaar. Deze geeft geen enkele informatie, ook niet in negatieve zin, over het kenmerk. Bij aftrekking van beide kansen wordt de formule:

$$a = \frac{f_a \cdot f_d - f_b \cdot f_c}{(f_a + f_c)(f_b + f_d)}$$

Bovenstaande formule is gelijk aan de formule voor de regressiecoëfficiënt van de operator op de klasse voor twee binaire variabelen. Het is dus mogelijk het lineaire regressiemodel uit de statistiek te gebruiken voor de bepaling van de waarde van de operator. Dit model geeft die lineaire functie tussen operator en klasse, waarbij de som van de kwadraten van de fouten minimaal is. In formule

$$t' = a \cdot T + c$$

waarbij T de bedoelde score van een element is,

t' de schatting van de score van de operator,

a de regressiecoëfficiënt, en

c een constante.

Het blijkt nu, dat als getracht wordt $(t-t')^2$, waarbij t de score van de operator is, voor de elementen zo klein mogelijk te houden a en c respectievelijk de volgende waarden hebben (zie o.a. Hoel, 1962, p.170 en 171):

$$a = \rho_{tT} \cdot \frac{\sigma_t}{\sigma_T} \quad \text{en} \quad c = \mu_t - a \cdot \mu_T$$

waarbij ρ_{tT} de populatiecorrelatie is van t en T,

σ_t, σ_T de populatie standaarddeviaties van resp. t en T, en

μ_t, μ_T de populatie gemiddelden van resp. t en T.

De regressieformule wordt dan:

$$(4) \quad t = a \cdot T + c + \delta$$

waarbij a en c gedefinieerd zijn met bovenstaande formules, en

$$\delta \text{ de meetfout is } t - t' = \delta$$

Formule 4 zal in paragraaf 3.4. verder worden uitgewerkt met behulp van het factoranalytische model (zie formule 5). Daarbij wordt de meetfout verder uitgewerkt. De constante c is alleen van belang bij absolute interpretaties van de scores (zie o.a. eind paragraaf 3.6.). Meestal worden scores relatief geïnterpreteerd, namelijk als intervalmaten met een arbitrair nulpunt.

Bij gebruik van standaardscores is c nul. In dat geval is een vergelijking van formule 4 met het klassieke model van de testtheorie mogelijk. Bij dit model wordt formule 4:

$$t = T + \delta$$

en is het populatiegemiddelde van δ nul en de populatiecorrelatie van T en δ ook nul (zie o.a. Lord en Novick, 1968, p.56, 3.1.1., 3.1.2. en 3.1.3.). Beide laatste aannamen gelden ook voor het lineaire regressiemodel en zijn uitwerking met behulp van factoranalyse. Het verschil is dat a de regressiecoëfficiënt van t (de geobserveerde score) op T (de ware score) één is (zie Lord en Novick, 1968, p.65, 3.7.1.).

Wanneer deze consequentie ($a=1$) wordt uitgewerkt voor het gegeven voorbeeld van twee binaire variabelen uit tabel 3.1.2., dan blijkt dat:

$$\frac{f_a \cdot f_d - f_b \cdot f_c}{(f_a + f_c)(f_b + f_d)} = 1$$

of

$$f_b(f_a + f_c) + f_c(f_d + f_b) = 0$$

Uit de calculus der volzinnen is ook af te leiden, dat de operator van een klasse een definitie is, in de zin van 'dan en dan alleen'. Alleen deze relatie geeft namelijk de mogelijkheid b en c als onwaar of fout te verklaren. Bij een implicatie zou bijvoorbeeld alleen mogelijkheid b onjuist zijn en is mogelijkheid c geen fout.

In de statistiek (zie o.a. Hoel, 1962, p.48) wordt mogelijkheid b 'type I fouten' genoemd of 'false positives' en mogelijkheid c 'type II fouten' of 'false negatives'. De waarde van de operator is nu te bepalen door na te gaan hoe groot de kans is op deze fouten bij het gebruik van de operator.

Dergelijke kansen kunnen worden geschat in een aselechte steekproef van elementen. Voor de eenvoud van de notatie zal allereerst hier worden uitgegaan van een eindige populatie van N elementen, waar mogelijkheid a; f_a maal optreedt, b; f_b maal, c; f_c maal en d; f_d maal. In tabel 3.1.2. staat de verdeling van deze frequenties. Op deze wijze is het voorlopig niet nodig een onderscheid te maken tussen populatie parameters en steekproefschattingen. Bovengenoemde kansen zijn dan werkelijk bekend. In de praktijk van het onderzoek is dit meestal niet het geval.

Tabel 3.1.2.

| | K(wa) | K(on) | |
|---------------------|---------------|---------------|-----------------------------|
| $\underline{O(wa)}$ | f_a | f_b | $(f_a + f_b)$ |
| $\underline{O(on)}$ | f_c | f_d | $(f_c + f_d)$ |
| | $(f_a + f_c)$ | $(f_b + f_d)$ | $N = f_a + f_b + f_c + f_d$ |

Bepaald kan nu worden de kans op een fout, wanneer een element, dat het kenmerk bezit, door de operator wordt geklassificeerd. Dit is een voorwaardelijke kans waarvan de formule als volgt luidt:

$$P \{O(on) | K(wa)\} = \frac{f_c}{(f_a + f_c)} \quad (f_a + f_c \neq 0)$$

De kans op een fout, wanneer het element het kenmerk niet bezit, is gelijk aan

$$P \{O(wa) | K(on)\} = \frac{f_b}{(f_b + f_d)} \quad (f_b + f_d \neq 0)$$

De kans dat de operator bij klassificatie van elementen een fout maakt is dus gelijk aan:

$$P \{O(on) | K(wa) \wedge O(wa) | K(on)\} = \frac{f_c \cdot f_b}{(f_a + f_c)(f_b + f_d)}$$

Bovenstaande kansen zijn alleen te berekenen wanneer $(f_a + f_c)$ en $(f_b + f_d)$ niet nul zijn, dat wil zeggen wanneer de klasse geen nul- of universele klasse is. Deze eis is al in paragraaf 2.4. gesteld voor doelmatige klassificaties.

Op overeenkomstige wijze kan de voorwaardelijke kans worden uitgerekend dat de operator geen fouten maakt. Deze kans is gelijk aan:

$$P \{O(wa) | K(wa) \wedge O(on) | K(on)\} = \frac{f_a \cdot f_d}{(f_a + f_c)(f_b + f_d)}$$

Beide kansen - de kans op fouten en de kans op geen fouten - zijn niet complementair, dat wil zeggen het is niet mogelijk bij kennis van de kans op fouten de kans op treffers te bepalen. Dit zou betekenen, dat de validiteit van een operator steeds met twee aparte indices zou moeten worden aangegeven. Het is nu nuttiger beide kansen te combineren tot één coëfficiënt. Hierbij geeft de aftrekking de meeste informatie.

Een operator namelijk waar de kans op fouten duidelijk groter is dan de kans op treffers geeft ook informatie en is, wanneer het coderingsvoorschrift wordt omgezet door $O=W$, te veranderen in $O \neq W$ zelfs een voorbeeld

Het model van de klassieke testtheorie is dus toepasbaar als

1. de kans op fouten nul is ($f_b = f_c = 0$) of,
2. de kans op fouten van de eerste soort nul is en de klasse een universele klasse is ($f_b = f_d + f_b = 0$) of,
3. de kans op fouten van de tweede soort nul is en de klasse een nul klasse is ($f_c = f_a + f_c = 0$).

Mogelijkheid 2 en 3 zijn al uitgesloten bij bespreking van de doelmatigheid van de klassificaties. Mogelijkheid 1 is zeer beperkt.

De conclusie is, dat de klassieke psychometrische betrouwbaarheidstheorie voor operatoren alleen bruikbaar is, wanneer de bedoelde klasse de operator zelf is, zoals het geval is bij onder andere een onderzoek van de inter-rater betrouwbaarheid.

Het gebruik van het lineaire regressiemodel stelt als voorwaarde, dat de gemiddelden van de verdelingen bij benadering op een rechte lijn liggen. Aan deze voorwaarde is bij een scatterdiagram van minstens één binaire variabele noodzakelijkerwijs altijd voldaan. Een rechte lijn wordt namelijk gedefinieerd door twee punten, dat wil zeggen door de twee gemiddelden van de twee verdelingen van de twee categorieën van de binaire variabele. Bij toepassing van de iteratieve clusteranalyse worden de operatoren altijd binair gecodeerd met één en nul en is dus bovenstaande voorwaarde geen probleem. Op theoretische gronden is het mogelijk een niet-rechthoekig verband tussen de klasse en de operator te veronderstellen. Door gebruik van verschillende coderingsvoorschriften voor eenzelfde operator kan de juistheid hiervan worden nagegaan, onder voorwaarde dat de operator in voldoende aantal (ordinale) eenheden is gekwantificeerd (zie paragraaf 4.3.).

Uit de formule van de regressiecoëfficiënt is af te lezen, dat de mate waarin de uitslagen van de operatoren zijn te voorspellen met behulp van de klassificatiescore niet alleen afhangt van de correlatie tussen operator en klassificatie, maar ook van de spreidingsmaten van beide variabelen. Uit deze formule is af te lezen, dat hoe groter de spreiding van de operator des te beter deze voorspelling is. De spreiding van de operator is van te voren te bepalen onder andere door de overeenstemming van de observatie met het verwachte antwoord ordinaal te coderen. Bij voldoende eenheden is het dan niet moeilijk de operator zodanig met nul en één te coderen, dat de spreiding maximaal is, namelijk een p/q verdeling van 0,50/0,50.

De correlatiecoëfficiënt geeft de validiteit van de operator onafhankelijk van de spreidingsmaten in de betreffende populatie. De correlatiecoëfficiënt heeft dus een algemenere betekenis. Om deze reden wordt in de clusteranalyse verder alleen maar gewerkt met correlatiecoëfficiënten. De keerzijde van de medaille is, dat voor gebruik van correlatiecoëfficiënten de steekproef van elementen voor het kenmerk k aselekt getrokken dienen te zijn. Toepassing van de analyse op extreme groepen is dan bijvoorbeeld niet toegestaan.

"Since least squares can be applied whether the x values (hier T scores) were fixed in advance or were obtained from random samples, the regression approach to studying the linear relationship between two variables is more flexible than the correlation approach. The interpretation of r as a measure of the strength of the linear relationship between two variables obviously does not apply if the values of x are selected as desired because the value of r will usually depend heavily on the choice of x values" (Hoel, 1962, p.172).

Aselecte trekking van de elementen is in de psychologie lang niet altijd uitvoerbaar. Toepassing van de analyse met regressiecoëfficiënten verdient dan zeker aanbeveling.

Moeilijkheden geeft ook de eis van homoscedasticiteit van het scatterdiagram. Aan deze eis dient beantwoord te worden, wanneer de populatiecorrelatie geschat wordt in een steekproef van elementen en de schattingsfout berekend moet worden. De correlatie van het regressiemodel is de z.g. productmoment correlatie, ook wel (Bravais-) Pearson correlatie genoemd. Bij een correlatie van een binaire en een continue variabele is dit de point-biserial r (zie o.a. Guilford, 1965). Das Gupta (1960) heeft erop gewezen, dat voor point-biserial's, als de populatie correlatie niet nul is, de voor de homoscedasticiteit noodzakelijke aanname, dat de verdelingen in de twee categorieën van de binaire variabele normaal verdeeld zijn met gelijke tweede en hogere momenten, niet juist is. De betrouwbaarheidsformules van de productmoment correlatie zijn dus niet bruikbaar voor de point-biserial r . In bijlage I.1. wordt verder op deze problematiek ingegaan. Daar zal blijken dat deze formules als benaderingsformules toegepast kunnen worden bij steekproeven groter dan 70 en p -waarden van de operator kleiner of gelijk aan 0,84 of groter of gelijk aan 0,16.

Het probleem van de homoscedasticiteit kan tot een bepaalde hoogte worden verminderd, wanneer gestreefd wordt naar operatoren met p/q verdelingen van 0,50/0,50. In dat geval is de aanname, dat de verdelingen van de twee categorieën voor het tweede moment gelijk zijn acceptabel, indien tenminste de elementen een symmetrisch-verdeelde T -score hebben. Daar bovenstaande al om andere redenen wordt nagestreefd, is het hier alleen een argument te meer voor een dergelijke verdeling van de operator.

Samenvattend kan worden gesteld dat primair de waarde van een operator gegeven wordt door zijn validiteit te bepalen door middel van de correlatie tussen operator en bedoelde klassificatie onder voorwaarde van aselechte trekking van de elementen. De psychometrische betrouwbaarheid van een operator blijkt verder alleen van belang wanneer getwijfeld wordt aan de objectiviteit van het coderings- of observatievoorschrift van de operator. Secundair zijn de problemen verbonden aan de uitvoering van de operator, die uiteraard ook mede zijn waarde bepalen.

Het is mogelijk de validiteit van de operator met andere indices dan de correlatie- of regressiecoëfficiënt aan te geven. Dergelijke indices zijn echter meestal ontwikkeld met behulp van een veel zwakkere theorie dan die van de lineaire regressie, waardoor zij wel een algemenere toepasbaarheid hebben echter vaak veel minder duidelijke antwoorden kunnen geven op gestelde vragen. Hierbij wordt onder andere gedacht aan de rank-biserial correlatie van Cureton (1956) en Bryden's (1960) coëfficiënt voor item-schaalbaarheid (een verdelingsvrije coëfficiënt). Bijnen (1968) geeft verder een overzicht van een groot aantal van dergelijke coëfficiënten. Daar het hier ontbreekt aan ervaring met deze indices, zullen er verder geen uitspraken over hun betekenis voor de bepaling van de waarden van een operator worden gedaan.

3.2. De bedoelde klassificatie

De meest toegepaste methode voor de constructie van een bedoelde klassificatie is het stellen van één operator als operationele definitie van de klassificatie. Deze werkwijze vindt vooral plaats in de experimentele- en sociale psychologie. Zo is voor Van de Geer (1957) de kannenproef van Luchins een operationele definitie van 'rigiditeit' en voor Mulder (1958, p.72) 'het zelf maken van de probleemoplossing' een operationele definitie voor 'zelfrealisatie'.

Eén operator is altijd een specificatie van één kenmerk, namelijk het kenmerk dat elementen bij dit observatievoorschrift niet of juist wel volgens de verwachting zich gedragen dan wel gedrag oproepen. Dit kenmerk is meestal vrij specifiek. Eigenschappen als 'rigiditeit' en 'zelfrealisatie' staan echter vrij hoog in de orde van de klassificatie. Terecht kan dus bij operationalisering met één operator van vrij abstracte begrippen worden afgevraagd of er dan niet sprake is van een partiële dekking en dus van een ongeoorloofde generalisatie. Is het niet correcter om te stellen dat het onderzoek van Van de Geer niet gaat over 'rigiditeit', maar over een bepaalde vorm van gedrag, zoals optreedt in de kannenproef en dat Mulder geen onderzoek doet naar 'zelfrealisatie' maar naar de betekenis om in een bepaalde opdrachtsituatie zelf het antwoord te kunnen geven in plaats van het te krijgen? Wanneer dit laatste het geval is, gaan beide onderzoeken over specifieke en dus weinig essentiële verschijnselen. In paragraaf 2.4. is dit al met de definitie van John Stuart Mill uiteengezet.

Een andere mogelijkheid is, dat de gekozen operator het basiskenmerk van de klasse psychometrisch zeer betrouwbaar bepaalt. In dat geval is één operator voldoende. Het bekende voorbeeld in deze is de thermometer als operator voor temperatuur. Het basiskenmerk van tempera-

tuur, waaruit alle andere kenmerken volgen, is de verhoging van de snelheid van de moleculen bij temperatuurstoename. Voor kwik betekent dit een vergroting van volume bij verhoging van temperatuur. Kwik is bijna niet te verontreinigen. Zijn uitzettingscoëfficiënt blijft dus bijzonder constant. Daarom is de kwikbuis een zeer goede operator voor temperatuur. Met veel minder zekerheid zou het zelfde gezegd kunnen worden van bijvoorbeeld de woordenlijst voor 'algemene intelligentie'. Deze genoemde mogelijkheid kan voor de kannenproef t.o.v. 'rigiditeit' en het zelf maken van de probleemoplossing t.o.v. 'zelfrealisatie' opgaan, is echter niet erg waarschijnlijk. Zowel Van de Geer als Mulder hebben echter geen pogingen ondernomen om dit aan te tonen. Een dergelijk onderzoek is evenwel uitvoerbaar wanneer

- a. de kenmerken van bovengenoemde begrippen worden geëxpliciteerd,
- b. deze kenmerken met behulp van operatoren worden gespecificeerd, en
- c. in een steekproef van elementen wordt nagegaan in hoeverre het mogelijk is de klassificaties van deze kenmerken te voorspellen met het basiskenmerk.

Uit bovenstaande volgt, dat de bedoelde klassificatie kan worden geconstrueerd door specificatie van de kenmerken van de eigenschap en gecontroleerd door na te gaan in hoeverre empirisch de geconstrueerde klassificatie voldoet aan de theoretisch gestelde verwachtingen.

Voor de constructie van tests zijn in de psychometrie een aantal methodieken ontwikkeld. Een test is een operationalisering van trekken of eigenschappen van individuen. Eigenschappen van stimuli kunnen op methodologisch overeenkomstige wijze worden geoperationariseerd. Er is in deze geen principieel verschil tussen beide groepen van elementen.

De psychometrie kent nu twee verschillende methodieken. De eerste is de zogenaamde interne validering (zie o.a. Guilford, 1954, p.417 t/m 433, Davis, 1966, en Lord en Novick, 1968, p.327 e.v.), waarbij 'items' worden geanalyseerd door hun moeilijkheidsgraad (verdeling van de itemscores) zowel als hun discriminatievermogen (correlatie met de totaalscore) in een steekproef van elementen te berekenen. De tweede methode is de zogenaamde externe validering. In het laatste geval wordt de validiteit van de 'items' bepaald door weer in een steekproef van elementen de correlaties van de items met een extern criterium te berekenen.

Beide methodieken worden in de literatuur niet duidelijk verantwoord. De eerste methode wordt meestal beschouwd als het homogen maken van de set van items ten gunste van de interne efficiency en de psychometrische betrouwbaarheid, de tweede dient om de predictieve validiteit van de set te verhogen.

'Items' vallen als zodanig onder de definitie van een operator, zoals gesteld in paragraaf 2.3. In dat geval zijn de totaalscore (bij de interne validering) en het externe criterium (bij externe validering) operationalisering van de bedoelde klassificatie.

Op die wijze beschouwd is er dan echter wel enige kritiek mogelijk op deze constructiemethoden. Externe validering kan alleen voldoen, wanneer het criterium een psychometrisch betrouwbare operationalisering is van het basiskenmerk. Vooralsnog moet worden vastgesteld, dat dergelijke operationalisering in de psychologie nog nauwelijks bestaan. Onderzoek van personeelsbeoordelingen (zie o.a. De Wolff, 1963) leert dat het criterium 'geschikt-ongeschikt' niet goed op deze wijze wordt geoperationaliseerd. Rapportcijfers (zie o.a. De Groot, 1966) voldoen ook niet als vastlegging van het criterium

'voldoende-onvoldoende'. Verder geven de oordelen van psychiaters duidelijk moeilijkheden bij bepaling van het criterium 'geestelijk gezond' (zie o.a. Stouffer, 1950, p.473 t/m 479). Genoegzaam is de problematiek bekend van de verhouding van dergelijke directe criteria ten opzichte van de uiteindelijke bedoelde. Het is bijzonder belangrijk verder onderzoek te verrichten om te komen tot goede operationalisering van eigenschappen als 'succes, gezondheid en geschiktheid'. Het is echter geen oplossing om ad hoc een bepaalbare klassificatie als *de* operationalisering van het basiskenmerk van deze criteria te beschouwen.

Ten aanzien van de interne validering merkt Hettema (1967) terecht op: "Via een homogeniteitsanalyse of itemanalyse wordt uitgemaakt welke items een homogeen universum representeren (correleren met de totaalscore) en deze items worden in de test opgenomen. Impliciet wordt daarmee het universum ingeëngd tot een homogene subset van het eerder gedefinieerde universum waardoor de representativiteit wordt prijsgegeven" (p.622). In de vorige paragraaf is er al op gewezen, dat wanneer de psychometrische betrouwbaarheid als criterium wordt gehanteerd voor een operator de bedoelde klassificatie de operator zelf is. Door nu zoveel mogelijk de operatoren van een test equivalent te houden, zoveel mogelijk zich te beperken tot een homogene set, wordt de psychometrische betrouwbaarheid van de set zo hoog mogelijk. Dit laatste is als volgt aan te tonen: operatoren zijn homogeen, wanneer zij herhalingen van zichzelf zijn, dat wil zeggen zo hoog mogelijk intercorreleren. Hoe hoger nu de gemiddelde intercorrelatie van de operatoren wordt, des te hoger zal de psychometrische betrouwbaarheid van de set van operatoren zijn. Dit is als volgt te bewijzen (hierbij wordt het bewijs gevolgd van Richardson, 1941, p.387 en 388). Uitgaande

van de lineaire relatie tussen ware en waargenomen score (zie paragraaf 3.1.) kan worden aangetoond (zie o.a. Gulliksen, 1950, p. 74 e.v. en Lord en Novick, 1968, p.112 e.v.), dat

$$r_{tt} = \frac{k \cdot r_{ii'}}{1 + (k-1) \cdot r_{ii'}}$$

Hierbij is r_{tt} de psychometrische betrouwbaarheid van de set van operatoren (de test),

$r_{ii'}$ de correlatie tussen twee equivalente parallelsets i en i' ,

k het aantal keer dat t groter is dan i .

Bovenstaande formule is de algemene Spearman-Brown formule voor testverlenging. Voor set i en i' kan nu operator i en i' worden gesubstitueerd. De set van operatoren heeft n operatoren, dus k wordt n . Als nu alle operatoren homogeen zijn vormen zij alle t.o.v. elkaar equivalente paralleltests met gelijke varianties en gelijke intercorrelaties. Voor $r_{ii'}$ kan dus de gemiddelde intercorrelatie \bar{r}_{ij} worden gesubstitueerd. De formule wordt dan:

$$r_{tt} = \frac{n \cdot \bar{r}_{ij}}{1 + (n-1) \bar{r}_{ij}}$$

of oplossend voor \bar{r}_{ij}

$$\bar{r}_{ij} = \frac{r_{tt}}{n - n r_{tt} + r_{tt}}$$

of

$$1/\bar{r}_{ij} = \frac{1}{r_{tt}} - n + 1 \quad (\bar{r}_{ij} \neq 0)$$

of

$$1/\bar{r}_{ij} - 1 = \frac{1}{r_{tt}} - n$$

Aan de laatste formule is af te lezen, dat hoe meer \bar{r}_{ij} naar één gaat, des te meer het linkergedeelte van de vergelijking naar nul gaat en des te meer n/r_{tt} gelijk moet worden aan n en dus r_{tt} één of maximaal wordt.

Onze conclusie is dus, dat wanneer met behulp van de methodiek van de interne validering de bedoelde klassificatie wordt geconstrueerd, alleen die operatoren voldoen, die onderling equivalent zijn. Hierdoor wordt echter de bedoelde klassificatie zeer specifiek en dus niet essentieel.

Bovenstaande problematiek is, weliswaar op een geheel andere wijze, onder woorden gebracht door Cronbach en Meehl (1955) met de theorie van de begripsvaliditeit ('construct validity'). Met deze theorie hebben zij getracht de eenzijdige definiering van validiteit door middel van externe criteria te doorbreken. De validiteit van een test wordt niet gegeven door de mate waarin deze test een bepaald criterium kan voorspellen, maar door de mate waarin de gepostuleerde eigenschap tot uitdrukking komt in de testscore. Naderhand heeft Cronbach samen met Rajartnam en Gleser (1963) deze kwestie verder uitgewerkt en wel met het begrip 'generaliseerbaarheid'. Het gaat om de mate waarin het mogelijk is met behulp van een geobserveerde score naar een universum van mogelijke scores te generaliseren. "Our rationale requires the investigator to start his study by defining the universe that interests him, and then to observe under two or more independently selected conditions within that universe. The calculations tell him how well the observed scores represent the universe scores. Since the universe is a construct that he introduces because he thinks it has explanatory or predictive power, an investigation of generalizability is seen to be an investigation of the 'construct validity' of the measure" (p.157). De schrijvers

gaan hierbij nog verder door te stellen dat met de theorie van de generaliseerbaarheid de theorie van de betrouwbaarheid en validiteit samenvallen. Een gezichtspunt, dat altijd door Tryon (1957) is verdedigd, die de betrouwbaarheidscoëfficiënt opvat als een maat voor de 'domain' validiteit van een test.

Deze laatste opvatting is onjuist, wanneer tenminste onder betrouwbaarheid de psychometrische betrouwbaarheid van de klassieke testtheorie wordt verstaan. Lord en Novick (1968, p.173 e.v.) hebben bovengenoemde theorie uitgewerkt voor de zogenaamde 'generic error of measurement' en wel met behulp van het variantie-analytisch model. Daarbij blijkt dat de 'generic' (afgeleid van 'genus' in tegenstelling tot 'specific' of 'species') meetfout altijd groter of gelijk is aan de meetfout van het klassieke psychometrische model (zie p.179, 8.3.9.).

In de hier ontwikkelde theorie en de theorie van de 'generic' ware score, ook wel de theorie van de onvolledige parallele maten genoemd, zijn duidelijk enige overeenkomsten aan te wijzen. Het universum van mogelijke metingen correspondeert met de set van gespecificeerde kenmerken van de eigenschap of construct. Het verschil is echter, dat hier een onderscheid wordt gemaakt tussen de set van kenmerken, die *nevenschikkend* uiteraard eindig is, en de set van mogelijke operatoren, die inderdaad een universum vormt. De operatoren dienen niet representatief te zijn voor dit universum, echter wel voor de geëxpliciteerde kenmerken. Om deze reden is het niet noodzakelijk dat de operatoren aselekt getrokken zijn uit het universum van mogelijke operatoren. De eis van aselechte trekking, die ten grondslag ligt aan de uitwerking met behulp van het variantie-analytische model, is in de praktijk betrekkelijk zelden te verwezenlijken. Om deze reden zal hier verder aan de theorie van de 'generic' ware score geen aandacht worden besteed.

De kenmerken van een bepaalde eigenschap kunnen geëxpliciteerd worden met behulp van de literatuur, zoals onder andere Hermans (1967, p.34) met acht kenmerken deed voor 'prestatiemotivatie'. Het is ook mogelijk hypothetisch een beschrijving te geven van het element (proefpersoon of stimulus), die volledig gekarakteriseerd wordt door de eigenschap, zoals onder andere Pennings (1966, p.56 e.v.) deed met 'taak-intrinsieke' en 'taak-extrinsieke werkmotivatie'. Wanneer een dergelijke omschrijving wordt gegeven in observeerbaar gedrag, geeft een verder representatieve specificatie weinig moeilijkheden.

Een probleem blijft, dat de meeste onderzoekers zich om praktische redenen vaak beperken tot één van de vele mogelijke specificaties van het kenmerk. Wanneer nu alle kenmerken, maar op één gelijksoortige wijze worden gespecificeerd, zoals onder andere bij het gebruik van vragenlijsten het geval is, wordt het mogelijk dat niet-bedoelde condities de uitslag van de set operatoren gaan bepalen.

Op dit laatste punt hebben vooral Campbell en Fiske (1956) en met hen Cook en Selltiz (1964) de aandacht gevestigd. Campbell en Fiske spreken over 'convergent and discriminant validation by the multitrait-multimethod matrix' en stellen voor systematische 'trekken' en 'methoden' te variëren. Iedere trek wordt met iedere methode bepaald. De overeenstemming tussen de methoden van één trek vormt de convergerende validiteit. Deze dient duidelijk beter te zijn dan die tussen de trekken van één methode; de discriminerende validiteit. De betekenis, die zij aan methoden hechten is vooral het gevolg van de vele bevindingen van methoden varianties in de vorm van antwoordinstellingen, apparatenfactoren en halo-factoren. Nu heeft iedere onderzoeker de neiging om de wijze waarop de operatoren geconstrueerd worden constant

te houden tot zelfs in details toe. Zo bestaat de PMT alleen uit Likert-schalen, de BIT alleen uit keuzen uit vier uitspraken, en de DAT, ISI en NDT alleen uit meer keuze items, waardoor bij deze laatstgenoemde intelligentietests de 'operation' van het divergerende denken bij geen enkele subtest aan de orde komt (zie Nederlands Instituut van Psychologen 1969, resp. p.207, p.166, p.87, p.62 en p.91). Een test, die veel beter volgens de principes van Campbell en Fiske is geconstrueerd, is onder andere de Strong Vocational Interest Blank (Ned. Inst. v. Psychol., 1969, p.165).

Het is dus begrijpelijk dat Campbell en Fiske de nadruk leggen op de noodzakelijkheid van methodenvariatie. Hun standpunt is echter als volgt algemeen te formuleren.

De uitslag van iedere operator dient bepaald te zijn door de relevante conditie van de bedoelde eigenschap. Daarnaast is het mogelijk, dat een aantal niet-relevante condities de uitslag beïnvloeden. Wanneer deze laatste condities ook constant in de set werkzaam zijn, invalideren zij de set van operatoren. Zo is het mogelijk dat om een hoge score te behalen in de DAT, ISI of NDT, niet de voorwaarde dat de proefpersoon intelligent is geldt (de bedoelde of relevante conditie), maar de voorwaarde dat de proefpersoon goed is in convergerend denken (een niet-relevante conditie).

Betrekkelijk weinig is nog bekend over de stimuli die in observatiesituaties het gedrag van de proefpersonen of beoordelaars beïnvloeden. De indruk is, dat het doel van het onderzoek, zoals gepercipieerd door de proefpersoon of beoordelaar, vaak een werkzaam aspect is. Zo blijkt de ABV en PMT (Van Oers, 1964, de Koning, 1968, Hermans, 1968) afgenomen in selectiesituaties (bij sollicitanten of

door selectie-instituten) andere normen te vereisen dan wanneer deze tests worden afgenomen in vrijblijvende situaties (door wetenschappelijke instellingen, voor experimentele doeleinden). Hofstee (1967) vind in absolute oordelen (Likert-schaal: zeer goed, goed, neutraal, slecht en erg slecht) plaats- en volgorde-effecten en wel in verschil met vergelijkende oordelen, waar ze niet optreden. In de laatste oordeelsvorm treden echter wel locatie-effecten op; bij twee alternatieven wordt het eerste gunstiger beoordeeld dan het tweede. Zijn conclusie is, dat bij vergelijkende oordelen beoordelaars de neiging hebben 'conservatievere' uitspraken te doen dan bij absolute oordelen. Boven is verder al een voorbeeld gegeven van een verwante psychologische eigenschap, die conditioneel kan zijn voor de uitslagen.

De relevante of bedoelde conditie wordt gegeven door de eigenschap, die moet worden geoperationaliseerd. Om die reden kan bijvoorbeeld een antwoordinstelling, die meestal niet-relevant is, een bedoelde conditie zijn, namelijk in die gevallen waar getracht wordt persoonlijkheidskenmerken door kwalitatieve prestatietests in de vorm van antwoordstijlen te bepalen (Frederiksen and Messick, 1959). Meestal zijn niet-relevante aspecten van de observatiesituatie: de vorm van de vraag of opdracht, de persoon van de proefleider en het tijdstip van de observatie binnen vantevoren-bepaalde, bij-het-observatievoorschrift-behorende, grenzen. In hoeverre deze te onderscheiden aspecten ook werkzame condities zijn is moeilijk te beantwoorden en hangt behalve van de aard van de specificatie ook af van de bedoelde eigenschap. Om die reden kan voorlopig het beste een algemeen voorschrift worden gegeven: varieer zoveel mogelijk bij specificatie van de kenmerken de niet-relevante aspecten.

Ieder aspect dat naast de bedoelde conditie constant wordt gehouden,

kan potentieel een (niet-relevante) conditie zijn.

Samenvattend kan worden gesteld, dat de bedoelde klassificatie kan worden geconstrueerd door explicatie van een set van nevenschikkende kenmerken van de bedoelde eigenschap en door een zodanige specificatie van deze kenmerken in operatoren, dat deze variëren in hun niet-relevante aspecten. Een dergelijke set van operatoren is theoretisch gezien representatief voor de bedoelde klassificatie. In een empirisch onderzoek zal moeten blijken in hoeverre de set van verwachtingen geverifieerd dan wel gefalsificeerd wordt. Belangrijk is hierbij om vast te stellen, dat de operatoren *niet* equivalent moeten zijn, zoals in de klassieke testtheorie wordt aangenomen. De uitslagen van de operatoren moeten alle bepaald worden door de relevante conditie, echter niet constant door niet-relevante.

3.3. De klassificatiemaat

Sinds Stevens' (1951, 1959) uiteenzettingen over de verschillende meetniveau's, is het gebruikelijk in de psychologie veel aandacht te besteden aan het probleem van de meting. De discussie heeft zich daarbij toegespitst op de vraag welke meetniveau's de verschillende statistische bewerkingen vereisen. Zo wordt het niet toelaatbaar geacht om het regressiemodel toe te passen op kwantificeringen die niet minstens een interval-niveau van meting hebben. Sinds de studie van Adams, Fagot en Robinson (1965) is het duidelijk geworden dat uitspraken over de toelaatbaarheid van bewerkingen van de kwantificeringen afhangt van de interpretaties, die aan de resultaten van de bewerkingen worden gegeven. Er zijn daarbij twee soorten van uitspraken te onderscheiden. Bij de eerste soort de z.g. 'unit-free statements' gaat het om de semantische betekenis van de uitspraken in relatie tot de meting, bij de tweede soort, de z.g. 'unit-specific statements', om de intrinsieke verifieerbaarheid. Zo is de vraag of de intelligentie van a gelijk is aan de intelligentie van b een semantische kwestie en de vraag of de testscore van a als overeenkomstig beschouwd kan worden met de testscore van b een intrinsiek-verifieerbare zaak.

Uit bovenstaande volgt, dat het noodzakelijk is aan de maten van de bedoelde klassificatie duidelijke betekenissen te geven, die theoretisch aanvaardbaar en voor wat betreft hun relatie met de observaties verifieerbaar zijn.

De pragmatische oplossing, die Lord en Novick (1986, p.21 e.v.) geven voor bovengenoemde problematiek, is in dit geval niet toepasbaar, omdat dan voor de bedoelde klassificatie een cirkelredenering zou worden toegepast. "If we construct a test score by counting

Bovengenoemde voordelen geven bij de meeste psychologische onderzoeken de doorslag ten gunste van nul en één coderingen. In hoeverre de analyse ook toepasbaar is bij andere soorten van coderingen is nog een open vraag. Een alternatieve scoring door middel van het gebruik van nominale gewichten, zoals o.a. bij de ABV plaats vindt (zie Ned. Inst. van Psychol., 1969, p.194) kan alleen worden overwogen na de empirische controle van de set. Apriori mag een dergelijke weging niet worden toegepast, en wel op grond van de eis van gelijke varianties, die voor de toepassing van de analyse geldt. Of een dergelijke kwantificering ook een betere benadering geeft van de bedoelde score is nog steeds niet duidelijk uitgemaakt. De psychometrische betrouwbaarheid wordt er wel door verhoogd (Davis and Fifer, 1959, Hofstee, 1963). Validiteitsverbeteringen zijn tot op heden echter niet gevonden. Nadelen zijn verder, dat de scoring tijdrovender is en de scores minder doorzichtig worden. De kwantificeringen zijn dan niet meer experimenteel onafhankelijk en een interpretatie van kansmaten is dus niet meer mogelijk. Het laatste geldt ook voor kwantificering van de operatoren in meer eenheden. Ter voorkoming van misverstand wijzen wij er nog op, dat het voor de dataverzameling aan te bevelen is om de operatoren te kwantificeren met zoveel mogelijk eenheden op minstens ordinaal meetniveau. In dat geval is het namelijk betrekkelijk eenvoudig om te voldoen aan het voorschrift van 50/50% p/q verdelingen.

Het informatieverlies van intervalmaten, die met nul en één worden gecodeerd, is in paragraaf 6.3.1. tot een bepaalde hoogte op zijn praktische relevantie bekeken. Vooruitlopend op dit onderzoek kan hier worden gezegd, dat de resultaten suggeren, dat bij steekproeven die duidelijk groter zijn dan honderd en/of bij een duidelijke

overpresentatie van de set (twintig of meer operatoren) de validiteit van de operatoren duidelijk beter worden geschat bij gebruik van de intervalmaten in verschil met hun nul en één coderingen. In alle andere gevallen schijnt het informatieverlies weinig tot geen betekenis te hebben. Wij merken hierbij op, dat het betrekkelijk zeldzaam is in psychologisch onderzoek om observaties als intervalmaten te verzamelen en waar dit plaats vindt bijvoorbeeld bij tijdsmaten zijn vaak psychologisch-theoretisch dergelijke maten hooguit als ordinale maten van psychologische eigenschappen te interpreteren.

up correct responses (zero-one scoring) and treating the resulting scale scores as having interval properties, the procedure may or may not produce a good predictor of some criterion" (p. 22). De voorafgaande paragraaf heeft voldoende duidelijk gemaakt, dat juist het ontbreken van (externe) criteria het noodzakelijk maakt dat de bedoelde klassificatie het criterium is.

Uitgaande van de gedachtengang dat elementen meer of minder gekarakteriseerd kunnen worden door de bedoelde eigenschap, is in deze studie als klassificatiemaat gekozen de kans op overeenkomst tussen het element en het ideaal-type van de bedoelde klassificatie.

Deze score is gelijk aan het aantal maal, dat het element volgens de operatoren in de klasse valt, gedeeld door het totaal aantal operatoren, waarmee het element is geklassificeerd. Het ideaal-type is de proefpersoon of de stimulus die volgens alle operatoren in de klasse valt. Het is een theoretische constructie, die niet (of zelden) in de werkelijkheid voorkomt. Met behulp van deze constructie is het echter mogelijk elementen te meten voor wat betreft de mate waarin zij in de bedoelde klasse vallen.

Bovengenoemde maat is als percentage van het aantal gestelde examen- of proefwerkvragen algemeen bekend. Hij is echter in onbruik geraakt omdat onvoldoende de representativiteit van de operatoren in acht werd genomen. Travers (1955) spreekt in dit verband over "the famous but discredited system of converting raw scores into percentages ..." en karakteriseert het als "... another attempt on the part of the teacher to convert a relatively meaningless score into one which has some commonly accepted meaning". Gardner (1962) zegt over deze scores het volgende: "The deficiencies inherent in these kinds of scores have

been discussed so many times in the literature no attempt will be made here to go into detail again. A few of the issues are the lack of comparability of per cent scores on the same test for different people, the lack of comparability from test to test, and the lack of algebraic utility" (p.7). Ebel (1962) komt echter met de 'content *standard* test score' en definieert deze score met "the per cent of a *systematic sample from a defined domain of tasks* which an individual has performed succesfully". Het verschil met de vorige score is de nadruk, die Ebel legt op de representativiteit.

Het begrip 'representatief' is in de vorige paragraaf uitgewerkt. Bovenstaande maat is dus als kans alleen goed interpreteerbaar als de operatoren een specificatie zijn van een set van nevenschikkende kenmerken. De operatoren dienen dus gelijkwaardig te zijn. Deze eis is intrinsiek-verifieerbaar en wel door empirisch te controleren of de operatoren overeenkomstige p-waarden en gelijke correlaties hebben met de bedoelde klassificatie. Bij dergelijke operatoren zijn dan de maten van verschillende sets van operatoren vergelijkbaar, evenals de maten van verschillende elementen bij dezelfde set.

Lord heeft in 1955 een overeenkomstige score, de 'proportion-correct' score, geïntroduceerd, naderhand door Keats en hem (1962) 'relatieve' score genoemd. Om die reden zal verder worden gesproken over relatieve score of klassificatiemaat, in onderscheid met bijvoorbeeld de standardscore.

De relatieve score wordt gemeten in het aantal eenheden dat de set aan operatoren telt. De nauwkeurigheid van de maat is dus een directe lineaire functie van het aantal operatoren. Hierdoor is het

mogelijk de maat als een benadering van een continue score voor te stellen. In de psychologie worden klassificaties meestal niet discreet, maar continue opgevat; personen zijn meer of minder intelligent, werkzaamheden meer of minder belastend. Hierdoor wordt het ook mogelijk de statistisch-mathematische theorieën voor continue variabelen op de scores toe te passen. Verder kan de maat worden geïnterpreteerd als de kans, dat een persoon zich gedraagt overeenkomstig personen met de bedoelde eigenschap of dat een stimulus gedrag oproept conform stimuli met de bedoelde kenmerken.

Een bezwaar van de maat is zijn begrensdheid. Kanswaarden lopen van nul tot en met één of van nul tot en met honderd procent. Het is bij deze maat niet mogelijk 'knapper' (voor individuen) of 'smakelijker' (voor etenswaren) te zijn dan het ideaaltype. Hierdoor moeten rekenkundige bewerkingen als de optelling en vermenigvuldiging op een andere wijze plaats vinden dan gebruikelijk. De relatieve score van twee sets van respectievelijk 8 en 10 operatoren is niet gelijk aan de som van de twee relatieve scores maar gelijk aan de uitslag van de twee sets samen onder voorwaarde dat de sets geen operatoren gemeenschappelijk hebben: 'mutually exclusive' zijn. Dus niet $6/8 + 5/10 = 5/4$, maar $\frac{6+5}{8+10} = 11/18$. De kanstheorie behandelt deze bewerkingsmogelijkheden in theorieën van de optelling en vermenigvuldiging. Het voordeel is echter wel dat bij dit uitgangspunt er sprake is van een fundamenteel meetsysteem met een duidelijk aan normen gebonden structuur, waardoor het mogelijk is na te gaan of rekenkundige en/of statistische bewerkingen van de maten toelaatbaar zijn (Adams, Fagot en Robinson, 1965).

Voor een goede interpretatie van de maat is het dus nodig dat

1. de operatoren worden gecodeerd met nul en één,
2. de operatoren experimenteel onafhankelijk zijn,
3. de p-waarden van de operatoren homogeen zijn,
4. de correlaties met de bedoelde klassificatie homogeen zijn.

De regels 3 en 4 zijn empirisch altijd maar tot een bepaalde hoogte bereikbaar. Hun betekenis hangt af van het doel van het onderzoek. Zo zal zuiver wetenschappelijk onderzoek veel strengere eisen stellen aan het meet-instrument dan dit realiseerbaar is bij toegepast onderzoek. Voorwaarde 2 is al gesteld bij de definitie van een operator (zie paragraaf 2.3.) en zal daarom niet verder als zodanig worden genoemd. De nauwkeurigheid is vervolgens afhankelijk van het aantal operatoren. Hoe meer operatoren des te nauwkeuriger de schaal.

Coderingen van operatoren met nul en één geeft informatie-verlies (zie formule van Shannon paragraaf 2.4.). Dit nadeel staat tegenover de volgende praktische en theoretische voordelen.

1. Bij nul en één codering van de operator, geeft de toepassing van het regressiemodel geen problemen ten aanzien van de aanname van de lineaire samenhang.
2. Er worden bij nul en één coderingen geen voorwaarden gesteld aan de data. Deze kunnen kwalitatief of kwantitatief van aard zijn, ook afwijkende verdelingen spelen geen rol. De analyse is voor wat dit punt betreft een 'non-metrische' of 'parameter-vrije' methode.
3. In één onderzoek kunnen met nul en één coderingen tegelijkertijd verschillende soorten van data worden geanalyseerd.
4. Grote aantallen variabelen in de orde van grootte van honderd tot vijfhonderd kunnen dan worden geanalyseerd.
5. De interpretatie van de score kan dan, zoals is aangetoond, op een

Per definitie is nu de somscore bij n operatoren

$$(6) \quad S_b = \sum_{i=1}^n z_{bi}$$

Interpolatie van formule 5 in 6 geeft:

$$(7) \quad S_b = F_{1b} \cdot \sum_i^n a_{i1} + F_{2b} \cdot \sum_i^n a_{i2} + \dots + F_{mb} \cdot \sum_i^n a_{im} + \sum_i^n a_i \cdot \delta_{bi}$$

Voor de toepassing van het centroid model bij correlaties moet S_b eerst worden gestandaardiseerd. Per definitie is de standardscore gelijk aan:

$$(8) \quad s_b = \frac{S_b - M}{\sigma_s}$$

Nu is M van S gelijk aan nul omdat S_b een som is van standardscores en σ_s gelijk aan de wortel uit $\sum_i^n \sum_j^n r_{ij}$, waarbij r_{ij} de intercorrelatie is tussen operator i en j . Wij krijgen dus na invulling van deze waarden in formule 8 en de waarde van S_b uit formule 7 in formule 8 dat

$$(9) \quad s_b = F_{1b} \cdot \frac{\sum_i a_{i1}}{\sqrt{\sum_i \sum_j r_{ij}}} + F_{2b} \cdot \frac{\sum_i a_{i2}}{\sqrt{\sum_i \sum_j r_{ij}}} + \dots + F_{mb} \cdot \frac{\sum_i a_{im}}{\sqrt{\sum_i \sum_j r_{ij}}} + \frac{\sum_i a_i \delta_{bi}}{\sqrt{\sum_i \sum_j r_{ij}}} \quad (\sum_i \sum_j r_{ij} \neq 0)$$

Wanneer nu het centroid model toepasbaar is op de factorstructuur van de intercorrelaties dan is per definitie (Thurstone, 1947, p.150 e.v. en Harman, 1960, p.193 e.v.)

$$(10) \quad (\sum_i a_{i1})^2 = \sum_i \sum_j r_{ij} \quad (\text{voorwaarde 2})$$

en

$$(11) \quad \sum_i a_{i2} = \sum_i a_{i3} = \dots = 0 \quad (\text{voorwaarde 3})$$

dus

$$(12) \quad s_b = F_{1b} + \frac{\sum_i a_i \cdot \delta_{bi}}{\sqrt{\sum_i \sum_j r_{ij}}}$$

De somscore is dan dus een schatting van de factorscore van de algemene factor. De meetfout waardoor s_b niet exact gelijk is aan F_{1b} wordt nu kleiner, des te groter de som van de intercorrelaties is (regel b) en des te groter het aantal operatoren (regel c). Het laatste is per definitie zo. De som van unieke meetfouten is in het universum van operatoren nul, verder is de correlatie tussen a en δ nul. Dus het product van beide gaat ook naar nul.

Worden de voorwaarden en regels samengevat, dan komen we tot het volgende overzicht.

Voorwaarden

1. De uitslagen van de operatoren moeten met nul en één worden gecodeerd.
2. De eerste centroïd van de intercorrelaties van de operatoren dient de algemene factor van de operatoren te zijn.
3. De groepsfactoren van de eventuele niet-relevante condities dienen in de set een ladingsom van nul te hebben.
4. De som van de intercorrelaties van de operatoren mag niet nul zijn.

Regels

- a. Zorg voor zoveel mogelijk 50/50% p/q verdelingen van de operatoren.
- b. Streef naar een zo groot mogelijke som van intercorrelaties van de operatoren.
- c. Zorg voor een zo groot mogelijk aantal operatoren.
- d. Streef naar zoveel mogelijk operatoren met gelijke correlaties met de score van de bedoelde klassificatie.

Hierbij is regel d een nadere uitwerking van de voorwaarde voor de interpretatie van klassificatiematen, zoals gesteld in paragraaf 3.3.

De correlaties met de ongewogen somscore zijn namelijk gelijk aan de clusterladingen onder voorwaarde dat de varianties van de operatoren overeenkomstig zijn. Dit bewijs wordt hier gegeven.

De lading in de eerste centroid a_{i1} is per definitie de clusterlading. Te bewijzen valt dus dat (zie Harman, 1960, p.194, 10.6)

$$(13) \quad a_{i1} = \frac{\sum_j r_{ij}}{\sqrt{\sum_i \sum_j r_{ij}}} = r_{is}$$

Nu is per definitie

$$(14) \quad r_{is} = \frac{\sum_b z_{bi} \cdot S_b}{N \cdot \sigma_i \cdot \sigma_s}$$

Daar σ_i gelijk is aan 1 (standaardscores) en $S_b = \sum_i z_{bi}$ (zie formule 6) wordt

$$(15) \quad r_{is} = 1/\sigma_s \cdot \sum_b z_{bi} \cdot (z_{b1} + z_{b2} + \dots + z_{bn}) \cdot 1/N$$

of

$$(16) \quad r_{is} = 1/\sigma_s \cdot \left\{ \sum_b z_{bi} \cdot z_{b1} + \sum_b z_{bi} \cdot z_{b2} + \dots + \sum_b z_{bi} \cdot z_{bn} \right\} \cdot 1/N$$

Nu is per definitie:

$$(17) \quad r_{ij} = \sum_b z_{bi} \cdot z_{bj} \cdot 1/N$$

en

$$(18) \quad \sigma_s = \sqrt{\sum_i \sum_j r_{ij}}$$

$$\text{Dus } r_{is} = \frac{\sum_j r_{ij}}{\sqrt{\sum_i \sum_j r_{ij}}} = a_{i1}$$

Dit bewijs geldt ook als $\sigma_i = \sigma_j$.

Bovengenoemde voorwaarden en regels vormen nu de constructievoorschriften voor de bedoelde klassificatie. Aan de voorwaarden moet beslist worden voldaan. De regels hebben niet dit stricte karakter. Verontachtzaming van de regels kan echter tot gevolg hebben dat de constructie

van de bedoelde klassificatie onvoldoende betrouwbaar plaats vindt. Het zijn optimaliseringsregels. Des te beter ze worden opgevolgd des te betrouwbaarder en nauwkeuriger is de constructie. Strikte uitvoering van de regels is praktisch echter onuitvoerbaar. Afhankelijk van de vereiste zekerheid die de onderzoeker wenst te bereiken, kunnen ze meer of minder stringent worden toegepast.

De iteratieve clusteranalyse beantwoordt nu als rekenmethode aan bovengenoemde voorschriften. Binnen de gestelde voorwaarden en regels zijn echter een groot aantal verschillende rekenmethoden mogelijk. Een beperkt aantal variaties zijn daarvan op niet-systematische wijze uitgeprobeerd. Het feit dat de clusteranalyse een toepassing is van de gestelde voorschriften mag dus niet worden opgevat als een bewijs dat deze analyse de best mogelijke is. Daarnaast blijft de mogelijkheid open dat de clusteranalyse impliciet een aantal regels toepast, die niet in de uitgewerkte theorie zijn onderkend.

3.5. Controle en exploratie

De twee belangrijkste voorwaarden die gesteld worden aan de constructie van de bedoelde klassificatie, zijn de eis van een algemene factor en van gebalanceerde groepsfactoren. Voortaan zal een dergelijke factorstructuur een cluster worden genoemd. De set van operatoren dient dus een cluster te vormen.

De constructie en dus ook de opzet van het onderzoek en de analyse moeten zodanig worden uitgevoerd, dat de hypothese van één cluster kan worden aangenomen of verworpen. De data van een steekproef van elementen dienen te beantwoorden aan de verwachting van één cluster in bovengenoemde zin. Hypothesen dienen toetsbaar te zijn, hetgeen inhoudt dat zij uitspraken toelaten over hun onwaar zijn, m.a.w. zij dienen falsifieerbaar te zijn (Fokkema, 1960, p.284).

Nu is iedere intercorrelatiematrix om te zetten in een centroïd factormatrix. Van deze eigenschap maakt Thurstone (1947) gebruik wanneer hij de centroïd methode als rekenkundig eenvoudiger extractiemethode dan de hoofdassenanalyse voorstelt. Iedere intercorrelatiematrix heeft een zwaartepunt. De juistheid van deze uitspraak is makkelijk vast te stellen. Intercorrelaties van variabelen kunnen immers worden voorgesteld met behulp van cosinussen van hoeken tussen vectoren in een $n(\text{operatoren}) + m(\text{factoren})$ dimensionale ruimte. Ieder op een dergelijke wijze grafisch voorgestelde intercorrelatiestructuur heeft noodzakelijkerwijs een zwaartepunt, waardoor de vector per definitie de centroïd is. Met de gebruikelijke factoranalytische technieken als de centroïd methode, diagonaal methode en hoofdassenanalyse, is het dus niet mogelijk de hypothese te falsificeren. Hierbij moet dan nog wel worden gewezen op het feit

dat de hoofdassenanalyse centroiden van de varianties geeft in plaats van de correlaties, zoals de centroid methode. Het verschil is in dit verband niet belangrijk - een correlatie in het kwadraat is een variantiemaat - en om die reden wordt er verder niet op ingegaan.

Het ontbreken van een mogelijkheid van falsifieerbaarheid dwingt tot het zoeken naar extra regels. Thurstone (1947) heeft deze gevonden met het principe van de 'simple structure', waaraan naderhand door onder andere Kaiser (1958) met het varimax criterium een objectieve vorm is gegeven. Dit criterium is om theoretische redenen niet bruikbaar. Het doel van dit criterium is na te gaan in hoeverre de intercorrelatiematrix past in de uni-factor oplossing, dat wil zeggen een oplossing van alleen groepsfactoren en geen algemene factor, zoals Harman (1960) opmerkt:

"In general, the rotation of axes in order to arrive at simple structure may be viewed as an attempt to reduce the complexity of the variables. The ultimate objective would be a uni-factor solution, in which each variable would be of complexity one" (p.113).

Er zal dus gezocht moeten worden naar andere regels.

De eerste regel, die hiervoor is geformuleerd, komt tot een bepaalde hoogte overeen met de gebruikelijke methode bij itemanalyse en luidt:

- de onderzoeker dient van tevoren aan te geven welke waarde (d.w.z. welke correlatie met de som- of clusterscore) nog als minimum acceptabel is.

Bij item-analyse worden items, die te laag correleren met de totaal-score van de set als onvoldoende beschouwd (zie Davis, 1965 en Guilford, 1954, p.417 e.v.). Het verschil met deze methoden is, dat in

de iteratieve clusteranalyse echter niet de beslissingen worden genomen op een ad hoc geconstrueerde totaalscore, maar op een clusterscore, die binnen de aanwezige mogelijkheden van de data voldoet aan de constructievoorschriften, zoals gesteld in de vorige paragraaf. Het is, zoals al naar voren gebracht, mogelijk voor iedere intercorrelatiematrix een cluster te construeren. Dergelijke clusters voldoen echter meestal betrekkelijk slecht, omdat een groot aantal operatoren vaak een lage tot zeer lage lading hebben in de eerste algemene factor en dus de som van de ladingen of item-totaal correlaties bij dit aantal items niet optimaal hoog is. Verder heeft het gebruik van operatoren met lage ladingen ook het bezwaar, dat dan ten onrechte de kenmerken die deze items vertegenwoordigen van toepassing geacht kunnen worden op de klassificatie. Door een minimumwaarde aan te geven definieert in feite de onderzoeker het cluster en wel afhankelijk van de mate waarin de waarde hoger of lager is; de onderzoeker strenger of minder streng is. Gaat het om de bergtoppen, eilanden of onderzeese gebergten? Deze analogie dringt zich op als de intercorrelaties van de operatoren worden voorgesteld als vectoren in een ruimte. De onderzoeker kan het cluster definieren als een meer of minder dichte vectorenbundel. Deze minimumwaarde operationaliseert de orde van de klassificatie. In paragraaf 2.5. is naar voren gebracht dat een dergelijke operationalisering mogelijk is met behulp van een overeenkomstindex. Bij een hoge waarde zullen veel, kleine en intercorrelerende clusters worden gevonden, die vrij specifiek zullen zijn. Bij een lage waarde weinig, grote en nagenoeg onafhankelijke clusters, die een algemenere betekenis zullen hebben.

De overeenkomst met het hiërarchische clusterschema van Johnson en Ward (zie tabel 2.5.1.) ligt voor de hand. Om deze reden zal de genoemde minimumwaarde voortaan de overeenkomstindex worden genoemd. Een belangrijk verschil tussen beide hiërarchieën blijft echter wel aanwezig. Terwijl bij Johnson en Ward een cluster in de orde van de klassificatie alleen groter of kleiner kan worden, is bij de iteratieve clusteranalyse afhankelijk uiteraard van de data, mogelijk dat in een lagere orde een cluster verdwijnt ten bate van twee grotere clusters. In paragraaf 5.3.2. (vergelijk tabel 2.5.1. met tabel 5.3.2.1.) wordt dit nader empirisch aangetoond. Denkelijk is dit het gevolg van de minder strenge definitie die in deze studie wordt gegeven aan een cluster. Clusters worden vaak opgevat als bestaande uit variabelen met een intercorrelatiematrix van de rang één, dat wil zeggen één algemene factor met n (variabelen) unieke factoren, dus zonder gebalanceerde groepsfactoren. Uit de vorige paragraaf is gebleken dat een dergelijke strenge definitie niet noodzakelijk is.

De tweede regel, die is geformuleerd om falsifieerbaarheid van de hypothese mogelijk te maken, is dat voor de niet-relevante condities adequate variabelen dan wel ad hoc geconstrueerde operatoren in het onderzoek dienen te worden opgenomen.

Bij éénzijdig geconstrueerde operatoren is het mogelijk een eerste centroïd te krijgen, die afwijkt van de bedoelde algemene factor. In feite is er dan sprake van een intercorrelatiematrix van twee algemene factoren waarin alle operatoren een positieve lading hebben of van sterke niet-gebalanceerde groepsfactoren. Hierdoor krijgt het zwaartepunt een afwijking in de richting van de groepsfactor of tweede algemene factor. De relevante conditie is niet alleen werkzaam in de uitslagen, maar ook een niet-relevante en wel op overeenkomstige wijze. Beide condities zijn in de operatoren

3.4. Constructievoorschriften

De bedoelde klassificatie kan worden geconstrueerd met behulp van een set van representatieve operatoren. Aan deze constructie zijn een aantal voorwaarden verbonden. Naar voren is al gebracht dat de uitslagen van de operatoren gecodeerd dienen te worden met nul en één. Terwijl ook al uitgebreid is gesproken over de p/q verdelingen van de operatoren. Vatten wij de voorafgaande beschouwingen samen dan zijn hiermee twee voorschriften geformuleerd en wel respectievelijk:

- codeer de uitslagen van de operatoren met nul en één, en
- zorg voor zoveel mogelijk 50/50% p/q verdelingen van de operatoren.

Het laatst genoemde voorschrift is niet alleen verbonden aan het gebruik van correlatiecoëfficiënten (zie paragraaf 3.1.), maar ook een waarborg voor de optimale informatie van een binaire klassificatie (zie paragraaf 2.4.), zowel als een mogelijkheid voor een interpretatie van de scores als klassificatiematen of relatieve scores (zie paragraaf 3.3.). Daarnaast berust de iteratieve clusteranalyse als rekenmethode op de aanname dat de varianties van de operatoren gelijk zijn. Een dergelijke voorwaarde maakt op het eerste gezicht een bijzonder strenge indruk. In een steekproef van elementen is het echter niet noodzakelijk dat p/q verdelingen exact hetzelfde zijn. Het gaat om de (onbekende) populatie-waarden. In paragraaf 4.3. wordt nagegaan welke consequenties het niet voldoen van de data aan deze voorwaarde heeft voor de rekenmethode. Daar zal blijken dat onder bepaalde omstandigheden afwijkingen geen ernstige gevolgen hebben. Praktisch blijkt de regel ook weinig moeilijkheden op te leveren, als de data maar in voldoende (ordinale) eenheden zijn verzameld en in een vooronderzoek is nagegaan of de verschillende onderscheidingen ook voorkomen.

In de voorafgaande theorie is het begrip 'representatief' nader uitgewerkt door te stellen dat een set operatoren representatief is als de relevante- of bedoelde conditie bij iedere operator de uitslagen mede bepaalt en de niet-relevante condities variabel of niet-constant optreden. Dit model is nog niet scherp genoeg geformuleerd. Deze theorie zal daarom eerst nader worden uitgewerkt.

De keuze is daarvoor gevallen op de theorie van de factoranalyse. De redenen van deze keuze zijn gedeeltelijk traditioneel en gedeeltelijk opportunistisch bepaald. Het is namelijk niet de bedoeling te pretenderen dat dit model isomorfer met psychische verschijnselen zou zijn dan eventuele alternatieve modellen. De mate van isomorfie van modellen berust veelal op een analogie-redenering en moet meestal blijken uit het succes van het model bij gebruik. Vooralsnog is er geen reden om aan dit succes voor het f.a. model te twijfelen. Burt (1941, 1949) Vernon (1950) en Eysenck (1947, 1951) beschouwen factoranalytische technieken als klassificatiemethoden. In dit opzicht conformeren wij ons aan de angelsaksische traditie in de psychologische methodenleer. De ontwikkeling van de psychologische testtheorie heeft daarnaast grotendeels plaats gevonden met behulp van het factoranalytische model. Zo gebruiken Cronbach, Rajaratnam en Gleser (1963, p.147) dit model voor hun theorie, terwijl Guilford (1954, 1959, 1964) en Cattell (1952) het model toepassen bij de constructie van tests. Daarenboven is er dankzij handboeken als die van Thurstone (1947) en Harman (1960) bijzonder veel bekend over deze theorie en zijn technieken. De genoemde overwegingen zijn echter niet doorslaggevend. De mogelijkheid blijft open, dat er betere modellen bestaan voor de geschetste problematiek.

gecorrleerd. Cronbach (1951) heeft aangetoond, dat groepsfactoren een geringe invloed hebben op de somscore in tegenstelling tot de invloed van een algemene factor, als het aantal variabelen of operatoren maar voldoende groot is. Dit geldt voor zwakke, niet-gebalanceerde groepsfactoren, bij sterke niet-gebalanceerde groepsfactoren helpt vergroting van het aantal operatoren betrekkelijk weinig en zeker niet als de bedoelde factor met de groepsfactor in de operatoren is gecorrleerd. Tegen vragenlijsten is niet ten onrechte de bedenking gemaakt dat algemene eigenschappen als 'neuroticisme' of 'prestatie-motivatie' alleen worden gediagnosticeerd door middel van verbaal gedrag in de vorm van zelfbeoordelingen. Het is niet onmogelijk dat hierdoor verbaal-intelligente proefpersonen worden bevoordeeld. Verificatie of falsificatie van een dergelijke verwachting is mogelijk door 'neuroticisme' of 'prestatie-motivatie' van vragenlijsten te vergelijken met de 'intelligentie' van de proefpersonen. Door bij de constructie van de bedoelde klassificatie de genoemde regel te volgen is de onderzoeker in staat dergelijke bedenkingen te weerleggen. De situatie is hierbij uiteraard hetzelfde als bij toetsing van iedere psychologische theorie: de constructeur moet bij de opzet van zijn onderzoek zelf zorg dragen dat controles op de geconstrueerde klassificatie kunnen worden uitgevoerd.

Beide regels kunnen niet beschouwd worden als de noodzakelijke oplossingen voor de onbepaaldheid van factoranalytische technieken. Daarvoor is in deze studie nog onvoldoende dit onbepaalde karakter geanalyseerd. De mogelijkheid blijft, dat ook de iteratieve clusteranalyse in dit opzicht niet scherp genoeg is. Daarnaast is het mogelijk, dat de analyse bepaalde regels toepast, die niet of onvoldoende zijn belicht.

Voorlopig kan worden gesteld dat de hypothese van één cluster op twee manieren kan worden gefalsificeerd, en wel

- a. doordat de operatoren een te lage correlatie hebben met de clusterscore, en
- b. doordat variabelen van niet-relevante condities correleren met de clusterscore.

In navolging van Campbell en Fiske (1959) is het eerste punt de toetsing van de convergerende validiteit genoemd en het tweede van de divergerende validiteit. Het is goed er hierbij op te wijzen, dat Campbell en Fiske, die spreken over "convergent and discriminant validation", deze onderscheidingen voor een enigszins ander doel gebruiken dan hier is gesteld.

De convergerende validiteit kan worden getoetst met de nul-hypothese dat alle intercorrelaties in de populatie nul zijn. Deze toetsing wordt in paragraaf 5.3.1 verder uitgewerkt.

De divergerende validiteit kan worden getoetst volgens de methode aangegeven in bijlage I.2. De correlaties dienen homogeen te zijn met als populatiecorrelatie de waarde nul of een waarde lager dan de gebruikte overeenkomstindex (zie ook hier paragraaf 5.3.1.).

Het gebruik van de analyse in bovengenoemde zin - door ons controlerend-toetsend genoemd (zie paragraaf 4.7.) - is betrekkelijk zeldzaam. Veel vaker wordt de analyse exploratief of controlerend-exploratief toegepast. Op een aantal gebieden is de theoretische kennis nog onvoldoende om de analyse toetsend te kunnen gebruiken.

Bij exploratief onderzoek tracht de onderzoeker in de door hem verzamelde data clusters te vinden, dat wil zeggen, sets van operatoren die beantwoorden aan het centroïd model. Dergelijke sets kunnen goede operationaliseringen van 'bedoelde' klassificaties zijn. Zij beantwoorden

aan de eis van de convergerende validiteit en zijn binnen de door de onderzoeker verzamelde data goed divergerend. Een zwak punt van exploratief onderzoek is de benoeming van de geconstrueerde klassificaties. Deze vindt ex post facto plaats en dient dus nader geverifieerd te worden. De Groot (1961) wijst erop dat exploratief onderzoek *vooronderzoek* is. "Wanneer het vervolg de exacte theorie- en/of hypothesevorming en toetsing, niet plaats vindt, is alles van weinig waarde" (p. 324). Een andere oplossing is dat de onderzoeker zich onthoudt van interpretaties en alleen de belangrijkste kenmerken van het cluster noemt. De in de paragrafen 5.2.1. en 5.2.4. vermelde onderzoeken zijn hiervan een voorbeeld.

Controlerend-exploratief onderzoek ligt tussen toetsing en exploratie in. Het onderzoek naar de convergerende validiteit wordt hierbij niet toetsend, maar explorerend uitgevoerd. Het voordeel is dat de onderzoeker dan duidelijker informatie krijgt over eventuele door hem over het hoofd geziene mogelijkheden. Deze vorm van onderzoek wordt vaak in de eerste fasen van schaalconstructies toegepast. Bij dergelijke onderzoeken worden vaak verschillende bedoelde klassificaties tegelijkertijd geconstrueerd. Daarom worden zij in dergelijke onderzoeken gebruikt voor controle van de divergerende validiteit. De schalen zijn dan ten opzichte van elkaar constructies van niet-relevante condities. De schalen dienen namelijk onafhankelijk ten opzichte van elkaar te zijn en wel ten behoeve van het principe van de spaarzaamheid ('parsimony').

3.6. De waarde van de geconstrueerde klassificatie

Meestal zal het wel mogelijk zijn een bedoelde klassificatie te construeren die aan de gestelde voorwaarden beantwoord. Desondanks blijft het mogelijk dat de constructie onvoldoende plaats heeft gevonden en wel omdat de gestelde regels slecht zijn vervuld. Het aantal operatoren dat ten aanzien van de convergerende validiteit voldoet kan te klein zijn. De som van de intercorrelaties van deze operatoren kan te laag zijn. De clusterladingen van de operatoren kunnen teveel uiteen lopen en/of de p-waarden van de operatoren kunnen sterk onderling verschillen.

De laatstgenoemde bevinding is al voor de analyse bekend. De mogelijke oplossingen voor deze problematiek worden behandeld in paragraaf 4.3. In 't algemeen kan worden gezegd, dat wanneer deze oplossingen niet voldoende zijn toegepast, clusters van dergelijke data, verkregen via de iteratieve clusteranalyse, geen of nauwelijks enige betekenis kan worden toegekend. Het zijn vaak kunstmatige clusters van operatoren met overeenkomstige p-waarden.

Wanneer de waarden van de operatoren teveel uiteenlopen, hetgeen blijkt uit de clusterladingen (correlaties met de somscore, te toetsen met de in bijlage I.2. gegeven procedure), dan is meestal met een te lage overeenkomstindex gewerkt. Een strengere grens geeft dan veelal een betere klassificatie.

De vervulling van de regels over het aantal operatoren en de som van de intercorrelaties kunnen met behulp van psychometrische betrouwbaarheidscoëfficiënten worden gewaardeerd. De psychometrische betrouwbaarheid, gedefinieerd als de correlatie tussen twee gelijkwaardige paralleltests, is namelijk een monotoon stijgende functie van de gemiddelde intercorrelatie van de operatoren, als van het aantal operatoren bij gelijk

— de som van de intercorrelaties van de operatoren mag niet nul zijn.

Verder zijn er nog twee extra regels te formuleren. Des te beter met deze regels rekening wordt gehouden, des te beter is de schatting van de score van de bedoelde klassificatie. Deze twee regels zijn:

- de som van de intercorrelaties van de operatoren dient zo groot mogelijk te zijn, en
- het aantal operatoren moet zo groot mogelijk zijn.

Allereerst zal nu worden aangetoond dat bovengenoemde voorwaarden en regels gelden op de wijze zoals naar voren is gebracht. Dit bewijs zal om redenen van duidelijkheid worden gegeven met behulp van standaardscores in plaats van ruwe- of relatieve scores. Dit is geen bezwaar omdat de iteratieve clusteranalyse alleen met correlaties werkt en correlaties van ruwe-, relatieve- en standaardscores exact dezelfde zijn.

Bij toepassing van het factoranalytische model wordt ervan uitgegaan dat (zie Harman, 1960, (2.8) p.13)

$$(5) \quad z_{bi} = a_{i1} \cdot F_{1b} + a_{i2} \cdot F_{2b} + \dots + a_{im} \cdot F_{mb} + a_i \cdot \delta_{bi}$$

Hierbij is z_{bi} de gestandaardiseerde uitslag van de operator i bij het element b ,

$a_{i1}, a_{i2}, \dots, a_{im}$ de factorladingen van de operator i in de m -factoren,

$F_{1b}, F_{2b}, \dots, F_{mb}$ de factorscores van het element b in de m -factoren,

$$a_i = \sqrt{1 - \sum_{i=1}^m a_{i.}^2}, \text{ en}$$

δ_{bi} de unieke meetfout.

Van de verschillende factoranalytische verbijzonderingen is het centroid model van Thurstone (1947), zoals ook beschreven door Harman (1960, p.193 e.v.) gekozen. Thurstone heeft dit model alleen om redenen van rekenkundige aard ontwikkeld en wel als compromis voor de hoofdassenmethode. "The centroid method of factoring and the centroid solution for the location of the reference axes are to be regarded as a computational compromise, in that they have been found to involve much less labor than the principal-axes solution" (Thurstone, 1947, p.178 en Harman, 1960, p.111). Het model heeft echter op zichzelf betekenis en wel als een bepaalde factorstructuuroplossing die isomorf is met de in paragraaf 3.2. ontwikkelde theorie over de bedoelde klassificatie. In dit model wordt er namelijk van uitgegaan, dat de intercorrelatiematrix van de operatoren adequaat beschreven kan worden met behulp van één algemene factor en een aantal groepsfactoren, waarvan de som der ladingen over de operatoren nul is. Indien de set van operatoren representatief is, moeten bij toepassing van de operatoren op een steekproef van elementen de data voldoen aan bovengenoemd model. Voor het 'variëren' van de niet-relevante condities dient dus nu 'balancing' te worden gelezen.

De somscore, zoals gedefinieerd in paragraaf 3.3., blijkt verder een schatting van de factorscore van de eerste centroid van de intercorrelaties te zijn, onder voorwaarde dat de som van de intercorrelaties niet nul is. Er zijn dus nog drie extra voorwaarden waaraan moet worden voldaan opdat de somscore ook inderdaad opgevat kan worden als een schatting van de score van de bedoelde klassificatie. Deze drie voorwaarden zijn:

- de eerste centroid van de intercorrelaties van de operatoren dient de algemene factor van de operatoren te zijn,
- de groepsfactoren van de eventuele niet-relevante condities dienen in de set een ladingsom van nul te hebben, en

blijven van de gemiddelde intercorrelatie. Het eerste punt is al aangetoond met het bewijs van Richardson (1941) en wel in paragraaf 3.2. Het tweede punt is als volgt te bewijzen. Bij deling van de algemene Spearman-Brown formule door n (het aantal operatoren) en substitutie voor r_{ii} van \bar{r}_{ij} (de gemiddelde intercorrelatie) wordt de betrouwbaarheid (r_{tt}):

$$r_{tt} = \frac{\bar{r}_{ij}}{1/n + \frac{n-1}{n} \cdot \bar{r}_{ij}} \quad (n \neq 0)$$

Hoe groter nu n (het aantal operatoren) wordt, des te meer gaat $1/n$ naar nul en $n-1/n$ naar één. De psychometrische betrouwbaarheid r_{tt} is dus een monotoon stijgende functie van n (het aantal operatoren) met als limietwaarde één.

De twee genoemde punten - de gemiddelde intercorrelatie en het aantal operatoren bij gelijkhouden van de gemiddelde intercorrelatie - zijn congruent met de in deze studie geformuleerde regels b en c (zie paragraaf 3.4.). Een onderzoek naar de psychometrische betrouwbaarheid, gedefinieerd met behulp van equivalente paralleltests, kan dus worden gebruikt om de waarde van de geconstrueerde klassificatie ten aanzien van deze twee regels aan te geven.

De psychometrische betrouwbaarheid heeft altijd een zeer belangrijk onderdeel gevormd van de testtheorie, zoals blijkt uit de overzichten van Wieggersma (1960) en Lord en Novick (1968, p.198 e.v.). Vergelijking van beide studies leert dat sinds 1960 er een cumulatieve ontwikkeling is geweest op dit terrein. Op het ogenblik zijn een zeer groot aantal verschillende bepalingsmogelijkheden, afhankelijk van de wijze waarop de ware- of foutenscore is gedefinieerd, beschikbaar.

Voor het hier in paragraaf 3.4. ontwikkelde model blijkt echter geen formule aanwezig. De psychometrische betrouwbaarheid van een cluster, gedefinieerd als een centroid structuur van intercorrelaties, zou het beste bepaald kunnen worden door berekening van de correlatie tussen de somscore en de factorscore uit formule 12. In dit geval ligt ook de meetfout vast. Aan deze mogelijkheid is echter in deze studie verder geen aandacht besteed.

Om voornamelijk praktische redenen zijn in deze studie de formules 20 en 21 van Kuder en Richardson (zie Wiegersma, 1960) gebruikt. Aan deze formules zijn een aantal nadelen, zowel als voordelen verbonden. De voordelen zijn dat de coëfficiënten snel en eenvoudig zijn te berekenen, de uitkomsten tot een bepaalde hoogte toetsbaar zijn (zie bijlage I.4.) en de formules correct zijn binnen de klassieke testtheorie. Met dit laatste punt zijn echter ook direct de nadelen van deze formules gegeven. In de klassieke testtheorie wordt de meetfout anders gedefinieerd dan in deze studie. Analyse van de verschillen en overeenkomsten tussen het factoranalytische model en het ware scoremodel leert (zie Lord en Novick, 1968, p.535 e.v. en Harman, 1960, p.15), dat de unieke meetfout van het factoranalytische model bestaat uit een specifieke meetfout en een random meetfout. Alleen wanneer nu de specifieke meetfout nul is, zijn de meetfouten van beide modellen gelijk. De specifieke factor wordt gegeven door die factor, welke ontstaat door de specificiteit van de set van operatoren en welke deze set *alleen* gemeenschappelijk heeft met een gelijkwaardige parallelset van operatoren. In deze studie vormt een dergelijke factor een foutencomponent. De parallelset dient niet gelijkwaardig, maar representatief te zijn (zie paragraaf 3.2.). Het is verder niet aannemelijk, dat een dergelijke specifieke factor geen invloed zou hebben op de somscore. Bij gebruik van de K.R. formules moet dus worden verwacht dat de uitkomsten een overschatting zullen geven van de werkelijke correlatie tussen somscore en bedoelde score.

Daarnaast hebben Novick en Lewis (1967, zie ook Lord en Novick, 1968, p.87 e.v.) aangetoond dat alleen als de twee parallel tests 'essentially true-score-equivalent' zijn coëfficiënt α gelijk is aan de psychometrische betrouwbaarheid van de test. In alle andere gevallen is deze coëfficiënt een 'lower bound', dat wil zeggen altijd lager dan de werkelijke waarde. Coëfficiënt α (zie Cronbach, 1951) is de algemene Kuder-Richardson formule. De K.R.(20) en (21) zijn verbijzonderingen van deze formule, namelijk voor die gevallen, waar de somscore een ongewogen optelling is van nul en één coderingen van experimenteel onafhankelijke operatoren. Bovenstaande geldt dus ook voor deze formules.

De voorwaarde van essentiële 'ware score' equivalentie - een noodzakelijke en voldoende conditie - houdt in, dat de ware score van element b bij operator i op een constante na gelijk is aan de ware score van element b bij operator j. In het model van paragraaf 3.4. wordt dit dan (afgezien van een andere ware score definiering):

$$a_{li} \cdot F_{lb} = a_{lj} \cdot F_{lb} + c$$

of

$$a_{li} = a_{lj}$$

Het gebruik van de formules stelt dus de vrij strikte eis, dat de clusterladingen (en de varianties) van de operatoren gelijk zijn naast de voorwaarde dat de specifieke meetfout nul is.

Het niet voldoen aan de regels a en d (zie paragraaf 3.4.), verlaagt derhalve de coëfficiënten ten opzichte van de bedoelde waarde; de verschillende meetfout definiering verhoogt de indices echter.

Het is dus mogelijk dat bovenstaande bedenkingen praktisch weinig relevant zijn. Om die reden is in één simulatie-onderzoek, namelijk nummer IV, nader nagegaan hoe de correlatie is tussen de ware score en de verkregen somscore ter vergelijking met de uitkomst van de K.R.(20). In dit onderzoek zijn data gesimuleerd voor 600 elementen bij 27 operatoren, die ieder een lading hebben in een algemene factor. Deze ladingen lopen van 0,40 t/m 0,80. De factorstructuur van dit onderzoek staat in bijlage III en de data in bijlage II.1.

De berekende productmoment correlatie tussen de ware score en de clusterscore van het eerste algemene cluster (ongewogen somscore van alle 27 operatoren) bleek nu 0,929 te zijn. De wortel uit de K.R.(20) gaf de waarde 0,944. Het verschil tussen beide uitkomsten is statistisch significant ($P \leq 0,003$). Geconcludeerd moet worden, dat de overschatting van de K.R. formule tengevolge van een andere definitie van de meetfout blijkbaar duidelijk groter is dan de onderschatting op grond van het niet voldoen aan de voorwaarde van de 'ware score' equivalentie.

Het gebruik van de K.R.formules voor de bepaling van de waarde van de geconstrueerde klassificatie is dus niet verantwoord. De uitkomsten van deze formules zijn overschattingen van de werkelijke waarden. Ontwikkeling van betere formules is dus vereist.

Rest nog het verschil aan te geven tussen de K.R.(20) en de K.R.(21) formule. Lord (1955, p.9) heeft aangetoond, dat het gebruik van de K.R.(21) juist is in die gevallen waar de onderzoeker de scores in absolute zin interpreteert, zoals het geval is bij schooltoetsen bij gebruik van tevoren vastgestelde 'cutting' scores. Het gaat om de meetfout, die gemaakt wordt bij de schatting van de c-waarde uit formule 4. Met deze meetfout wordt door de K.R.(21) wel, door de K.R.(20) geen rekening gehouden. De uitkomst van de K.R.(21) is daarom altijd lager dan die van de K.R.(20).

4. Uitvoering van de analyse

4.1. De steekproef

De samenstelling van de steekproef wordt in de eerste plaats bepaald door de doelstelling van het onderzoek. Daarom zal hier een onderscheid worden gemaakt tussen onderzoeken die *de klassificatie* van een aantal heterogene elementen beogen en die waarvan het doel *schaal- of testconstructie* is. De eerstgenoemde onderzoeken vereisen een ander gebruik van de analyse dan de tweede. Ook bij de vorming van de steekproef moet voor de eerste doelstelling op andere zaken worden gelet dan bij de tweede.

Voor *klassificatiedoeleinden* is de representativiteit van de steekproef van elementen een belangrijke kwestie, zoveel mogelijk moet in dat geval getracht worden de steekproef aselekt of stratified te trekken. Met afwijkingen bij de vorming van de steekproef moet bij de interpretatie van de resultaten rekening worden gehouden.

Voor *schaal- of testconstructie* is de belangrijkste voorwaarde, dat de elementen heterogeen zijn in de eigenschap, waarvoor schalen of tests worden geconstrueerd. De elementen moeten verschillen in de mate waarin ze behoren tot de bedoelde klassificaties. Het is niet mogelijk de operatoren te controleren als niet aan deze voorwaarde is voldaan. Binnen de mogelijkheden van steekproeftrekking is het verder belangrijk macro-sociale variabelen te gebruiken als geslacht, leeftijd of opleiding om tot een duidelijk te omschrijven populatie te komen, waarvoor de steekproef representatief is. Een dergelijke steekproef is dan een wel omschreven referentie- of normgroep, die aangeeft voor welke elementen de test of schaal is gecon-

strueerd. Het is ook mogelijk te werken met extreme groepen van elementen in de bedoelde constructs. In dit geval kan de analyse echter niet met correlatiecoëfficiënten worden uitgevoerd (zie paragraaf 2.1.).

Een tweede punt waar de onderzoeker voor wordt geplaatst, is de optimale grootte van de steekproef. De beslissing hierover wordt bepaald door de kosten verbonden aan de trekking van één element en de informatietoename per element. Het is algemeen bekend dat de informatie, die ieder extra-element oplevert, afneemt. Hoe groter de steekproef wordt, des te minder informatie geeft de trekking van een extra-element. Omdat de kosten verbonden aan de trekking van een element veelal gelijk blijven is het daarom mogelijk te spreken over een optimum. In de psychologie wordt - meer op grond van traditie dan op grond van berekeningen - dit optimum meestal gevonden tussen de honderd en vijftienhonderd elementen en wel afhankelijk van het doel van het onderzoek. Onderzoeken bij meer dan vijftienhonderd elementen zijn betrekkelijk zeldzaam. Als minimumgrootten gelden steekproeven van vijftientwintig tot veertig elementen. Steekproeven kleiner dan deze minimumwaarden worden alleen getolereerd wanneer de kosten verbonden aan de trekking van één element zeer hoog zijn. In de klinische psychologie is een dergelijke situatie vaak aan de orde.

De betrouwbaarheid van de analyse wordt bepaald door de grootte van de steekproef. Hoe groter de steekproef, des te kleiner de schattingsfouten zijn van de correlaties. De schattingsfout is ook een functie van de hoogte van de correlatie. Hoe lager de correlatie, des te groter de mogelijke fout en des te groter dus de optimale steekproef moet zijn.

De voorafgaande uiteenzettingen zijn nader geconcretiseerd in een onderzoek. Het ging hier om de antwoorden van 720 zusters op de voor de proefgroep aangepaste N.I.P.G. Satisfactieschaal (Wiegersma en Gemert, 1961). De gegevens van dit onderzoek staan in bijlage II.2. De schaal bestaat uit 15 uitspraken over het werk en wordt beantwoord met ja, geen mening of nee, inzoverre volgens de betrokkene de uitspraak van toepassing is op zijn/haar werksituatie. De verwachting was, dat de vijftien operatoren één cluster zouden vormen.

Toepassing van de clusteranalyse (controlerend-exploratief) gaf inderdaad één cluster. In tabel 4.1.1. staat het resultaat. De overeenkomstindex die is gebruikt is 0,30.

Tabel 4.1.1.

| Resultaten onderzoek Satisfactieschaal | | | % niet |
|----------------------------------------------------------------------------------------------|----------|------|------------|
| Vragen | ladingen | p | beantwoord |
| 1. De plaats waar ik sta/zit te werken bevalt mij wel. | 0,61 | 0,16 | 0,1 |
| 2. Mijn werk maakt me gejaagd en zenuwachtig. | -0,59 | 0,25 | 1,0 |
| 3. Mijn superieur kan goed met zijn/haar ondergeschikten opschieten. | 0,55 | 0,35 | 2,7 |
| 4. Mijn werk is te vermoeiend. | -0,61 | 0,28 | 1,0 |
| 5. Ik zou hier graag ander werk hebben. | -0,65 | 0,22 | 0,3 |
| 6. Dit is een prettige instelling/school om te werken. | 0,62 | 0,31 | 6,1 |
| 7. Deze instelling/school wordt goed geleid. | 0,59 | 0,39 | 6,3 |
| 8. Ik vind mijn werk te eentonig. | -0,63 | 0,16 | 1,3 |
| 9. Er zijn onder mijn naaste medewerkers veel onderlinge wrijvingen en strubbelingen. | -0,58 | 0,36 | 0,9 |
| 10. In mijn community vinden ze dat ik goed werk heb. | 0,48 | 0,29 | 1,0 |
| 11. Aan de bij het werk noodzakelijke hulpmiddelen wordt hier veel zorg en aandacht besteed. | 0,46 | 0,17 | 0,7 |
| 12. Ik voel me aan het einde van de werkdag meestal voldoende fit. | 0,48 | 0,49 | 0,8 |
| 13. Ik werk in een prettige omgeving. | 0,67 | 0,27 | 0,4 |
| 14. Daar waar ik werk is het altijd rommelig. | -0,55 | 0,19 | 0,9 |
| 15. Ik heb plezier in mijn werk. | 0,62 | 0,16 | 0,3 |

K.R.(21) = 0,84

Door aselechte trekking werden nu twee groepen van 120, één van 100, twee van 80, twee van 60 en drie van 40 proefpersonen gevormd. Op deze groepen werd de analyse herhaald met dezelfde overeenkomstindex van 0,30. In tabel 4.1.2. staan de resultaten van de analyse van deze 10 groepen voor de eerste cluster.

Tabel 4.1.2.

Resultaten 10 groepen

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----------|---------|-------|-------|-------|-------------|-------|-------------|--------------|-------------|--------------|--------------|
| | N = 720 | 120 | 120 | 100 | 80 | 80 | 60 | 60 | 40 | 40 | 40 |
| 1 | 0,61 | 0,56 | 0,66 | 0,66 | 0,60 | 0,45 | 0,66 | 0,79 | 0,51 | 0,71 | 0,61 |
| 2 | -0,59 | -0,55 | -0,67 | -0,53 | -0,54 | -0,66 | -0,52 | -0,52 | -0,74 | -0,60 | -0,70 |
| 3 | 0,55 | 0,57 | 0,43 | 0,51 | 0,59 | 0,56 | 0,60 | 0,47 | 0,63 | 0,79 | <u>0,13</u> |
| 4 | -0,61 | -0,55 | -0,61 | -0,64 | -0,61 | -0,65 | -0,51 | -0,59 | -0,71 | -0,56 | -0,61 |
| 5 | -0,65 | -0,58 | -0,71 | -0,63 | -0,63 | -0,63 | -0,51 | -0,82 | -0,68 | -0,57 | -0,81 |
| 6 | 0,62 | 0,67 | 0,54 | 0,59 | 0,64 | 0,62 | 0,59 | 0,77 | 0,71 | 0,64 | 0,64 |
| 7 | 0,59 | 0,74 | 0,58 | 0,52 | 0,57 | 0,53 | 0,55 | 0,65 | 0,63 | 0,57 | 0,63 |
| 8 | -0,63 | -0,55 | -0,65 | -0,58 | -0,57 | -0,56 | -0,62 | -0,69 | -0,75 | -0,80 | -0,69 |
| 9 | -0,58 | -0,61 | -0,64 | -0,68 | -0,45 | -0,52 | -0,53 | -0,51 | -0,71 | -0,70 | -0,50 |
| 10 | 0,48 | 0,45 | 0,43 | 0,60 | 0,54 | 0,43 | <u>0,30</u> | 0,56 | <u>0,07</u> | 0,51 | <u>0,24</u> |
| 11 | 0,46 | 0,56 | 0,39 | 0,51 | <u>0,26</u> | 0,54 | 0,62 | 0,47 | 0,45 | <u>0,26</u> | <u>0,16</u> |
| 12 | 0,48 | 0,51 | 0,49 | 0,49 | 0,53 | 0,63 | 0,47 | <u>0,06</u> | 0,50 | 0,53 | 0,57 |
| 13 | 0,67 | 0,62 | 0,71 | 0,74 | 0,71 | 0,50 | 0,68 | 0,81 | 0,49 | 0,58 | 0,74 |
| 14 | -0,55 | -0,55 | -0,67 | -0,55 | -0,60 | -0,63 | -0,66 | <u>-0,28</u> | -0,59 | <u>-0,21</u> | <u>-0,28</u> |
| 15 | 0,62 | 0,55 | 0,64 | 0,67 | 0,46 | 0,55 | 0,66 | 0,85 | 0,47 | 0,69 | 0,75 |
| K.R. (21) | 0,84 | 0,84 | 0,84 | 0,86 | 0,82 | 0,83 | 0,82 | 0,88 | 0,86 | 0,86 | 0,85 |

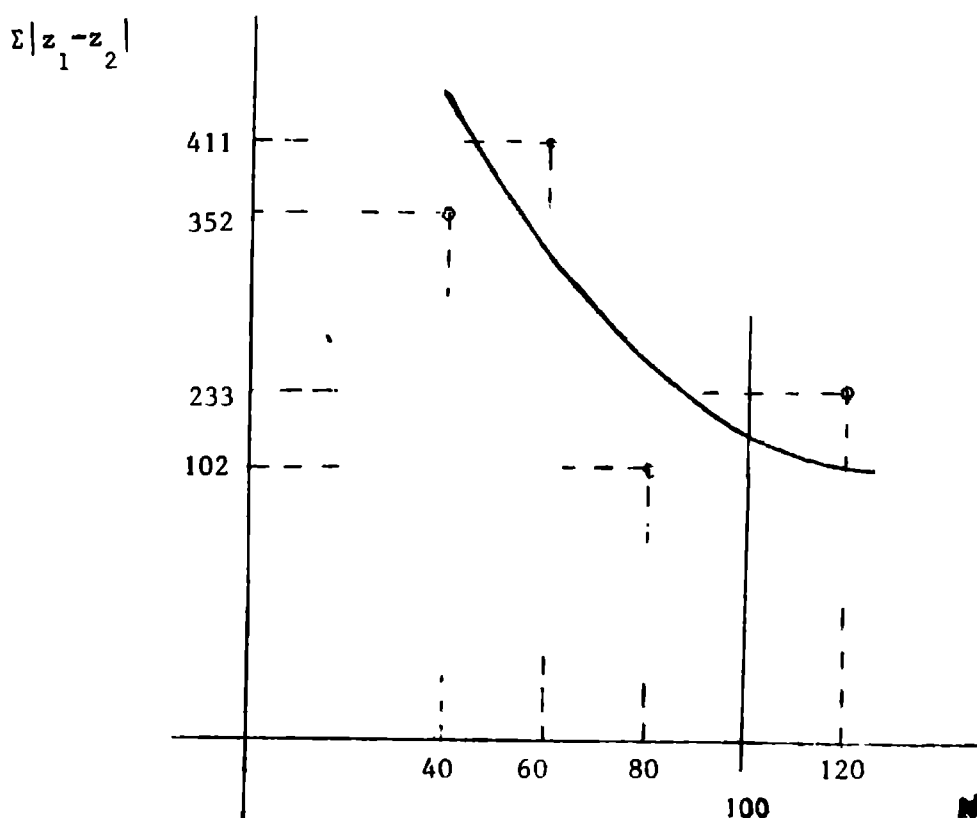
De onderstreepte waarden zijn de correlaties van de variabelen, die niet opgenomen zijn in het cluster. Deze uitval neemt toe bij verkleining van de steekproefgrootte en vindt vooral plaats bij variabelen met lage ladingen: 11, 10, 12, 14 en 3. Bij de laatste groep vormen variabele 14 en 11 een apart cluster met een K.R.(21) van 0,70. Deze waarde is niet sig-

nificant voor 1%, echter wel voor 5%. De grens is in het eerste geval 0,76 en bij 5% 0,61 (zie bijlage I.4.). De overige variabelen vormen allen een specifiek cluster, dat wil zeggen een cluster dat uit één variabele bestaat. Wij wijzen ook nog op de betrekkelijk grote stabiliteit van de K.R.(21) in vergelijking met de ladingen. Het gebruik van de K.R.(21) formule is in dit geval minder juist. Beter is de berekening van de K.R.(20). Daar de K.R.(21) echter de normale output is van het gebruikte computerprogramma, zijn de berekende uitkomsten ter algemene informatie vermeld.

De betekenis van de steekproefgrootte is in figuur 4.1.1. nader aangegeven. In deze figuur is de som van de absolute verschillen in z -waarden tussen de groepen met een gelijke N uitgezet voor de verschillende steekproefgrootten; 120, 80, 60 en 40.

Figuur 4.1.1.

Som absolute z -verschillen tussen groepen met gelijke N uitgezet voor de grootte van de groepen.



Uit figuur 4.1.1. is af te lezen dat het de moeite loont om bij een gemiddelde intercorrelatie van 0,35 de steekproefgrootte zoveel mogelijk op minstens honderd te brengen. De nauwkeurigheid neemt tot honderd progressief toe, daarna is de winst bij eenzelfde vergroting van de steekproef geleidelijk aan minder. Tevens kan uit dit onderzoek worden geconcludeerd, dat specifieke clusters (en clusters met K.R. waarden beneden de 1% significantie-drempel) weinig betekenis hebben.

4.2. De operatoren

Bij de keuze van de operatoren is er een verschil te maken tussen exploratief- en controle-onderzoek. In het laatste geval alleen construeert de onderzoeker de operatoren als specificaties van een kenmerk met zoveel mogelijk variatie in de niet-relevante aspecten (zie paragraaf 3.2.). Meestal echter zijn zonder veel vooronderstellingen de data van een aantal variabelen systematisch van elementen verzameld en wordt naar een doelmatige en overzichtelijke beschrijving van deze data gezocht. Door de afwezigheid van een duidelijk uitgangspunt bij de keuze van de operatoren zijn dergelijke onderzoeken exploratief.

De ervaring is, dat de meeste onderzoekers ook bij exploratief onderzoek, zich op researchstrategische gronden beperken in het aantal variabelen. Eerst worden een aantal schalen geconstrueerd en in een tweede bewerking of onderzoek wordt dan nagegaan hoe de relaties van de schaal scores zijn met externe criteria, macro-sociale variabelen of andere variabelen die geen items zijn. Deze researchstrategie is niet erg efficient, beter is het om direct al deze variabelen gezamenlijk te analyseren en de analyse dan controlerend-exploratief toe te passen, waarbij alleen wordt nagegaan welke clusters met de items zijn te construeren. De items vormen dan een van tevoren opgegeven set of pivot. In dat geval kan direct worden bepaald welke clusters behalve qua convergerende validiteit voldoen, ook qua divergerende of predictieve validiteit betekenis hebben. Het is mogelijk dat een klein, minder betrouwbaar cluster predictief belangrijk is. Bij de eerstgenoemde strategie zouden dergelijke 'verstopte' clusters zijn verwaarloosd. Verder is dan de interpretatie van de clusters eenvoudiger en wordt

de onderzoeker geattendeerd op relaties met andere gegevens, die hij anders misschien over het hoofd zou zien. Het is verder geen bezwaar als de varianties van de extra variabelen afwijken van de gemiddelde variantie, omdat de analyse alleen op de opgegeven operatoren wordt uitgevoerd. Wel dienen ze gecodeerd te worden met nul en één. Omdat ze niet opgenomen worden in de pivot is experimentele onafhankelijkheid dan ook geen voorwaarde.

Ook is het mogelijk op deze wijze omgekeerd te werk te gaan door namelijk eerst ieder criterium apart als pivot op te geven en dan in een tweede bewerking pivots van correlerende items controlerend-exploratief te laten analyseren (zie paragraaf 5.3.3.).

Het sterke punt van de iteratieve clusteranalyse is namelijk, dat deze methode, zelfs met alleen een rekenmachine, grote aantallen variabelen economisch verantwoord kan analyseren. Het is niet nodig bij het gebruik van deze analyse om bepaalde gegevens weg te laten, zoals bijvoorbeeld Wassenaar (1956) moest doen: "Omdat dit prakties onmoontlik was om al die gegewens gelijktijdig deur middel van faktorontleding te verwerk, is 24 variante vir die doel só gekies dat elke soort toets verteenwoordig is" (p.47). Genoegzaam is in het vorige hoofdstuk erop gewezen, dat het aantal operatoren zo groot mogelijk moet zijn. Dit punt vindt ook zijn bevestiging in de simulatie-onderzoekingen (zie paragraaf 6.3.1.). Klassificatie van gegevens is verder vooral van belang als er veel variabelen zijn. Klassificatiemethoden, die dus grote aantallen variabelen om praktische redenen niet kunnen analyseren, hebben daarom een beperkte betekenis. Er zijn weinig redelijke argumenten te geven voor de werkwijze, die door onder andere Wassenaar wordt gevolgd. Nog steeds zijn statistische technieken hulpmiddelen van de psychologie en niet

omgekeerd. Bij gebruik van de clusteranalyse moet het voordeel van het grote aantal variabelen worden gebruikt, dus geen testscores, als operatoren opnemen (met uitzondering van genoegzaam gevalideerde tests) maar items, en geen verzamelde data weglaten, omdat ze niet interessant zouden zijn. Dit laatste moet juist blijken en waarom zijn ze dan verzameld?

Operatoren als de persoon van de proefleider, het tijdstip van de dataverzameling, de plaats van de verzameling, dienen feitelijk in ieder onderzoek als routinevariabelen te worden opgenomen voor een controle op de divergerende validiteit. Het bezwaar, dat dergelijke gegevens kwalitatief van aard zijn, geldt namelijk niet voor de clusteranalyse.

De onderzoeker, die controlerend te werk wil gaan, wordt aangeraden vooral de paragrafen 3.2. en 3.5. door te lezen. De conclusies van deze paragrafen luiden in het kort als volgt.

1. Omschrijf de eigenschappen van de proefpersonen of stimuli, die moeten worden geoperationaliseerd.
2. Expliciteer de belangrijkste nevenschikkende kenmerken van deze eigenschappen en/of geef een beschrijving van de persoon of stimulus, die volledig door de eigenschap wordt gekarakteriseerd, in observeerbaar gedrag.
3. Specificeer de kenmerken met behulp van observatie- en coderingsvoorschriften en wel zodanig, dat niet-relevante aspecten gevarieerd zijn.
4. Doe hetzelfde (stappen 1 t/m 3) voor al die condities, waarvan verwacht mag worden dat zij de constructie kunnen invalideren en/of kies valide variabelen voor deze niet-relevante condities.

Variaties op bovengenoemde werkwijze wordt verder aan de onderzoeker overgelaten. In dit verband wordt nog gewezen op de mogelijkheid om een hiërarchisch-structuuronderzoek te doen door niet alleen nevenschikkende kenmerken, maar ook boven- en onderschikkende kenmerken te expliciteren (zie ook paragraaf 5.3.2.).

Hierbij dient te worden opgemerkt, dat voor ieder kenmerk minstens zes operatoren dienen te worden gespecificeerd. Bij een kleiner aantal is de kans te groot dat de constructie van de bedoelde klassificatie mislukt. Het is daarbij niet nodig, dat de gespecificeerde operatoren volledig en altijd zullen voldoen. De theorie van het vorige hoofdstuk en dus ook de analyse gaat ervan uit, dat bij gebruik van de operatoren toevallige- en constante fouten optreden.

4.3. Omzetting van de data in een nul en één matrix

De eerste voorwaarde en regel voor de analyse luiden (zie paragraaf 3.4.): "de data coderen in één en nul met bij benadering gelijke p-waarden van 50%". Wanneer de observaties verzameld zijn in zoveel mogelijk klassen op ordinaal meetniveau, is bovengenoemd constructievoorschrift vrij eenvoudig uit te voeren met behulp van bijvoorbeeld de mediaan. In tabel 4.3.1. staat een gefingeerde matrix van nul en één coderingen. Tussen haken wordt daarbij de oorspronkelijke waarde vermeld.

Tabel 4.3.1.

Matrix van nul en één coderingen
met coderingsvoorschriften

| | operatoren | | | | |
|-----------------------------|-------------------------------------------------------------------------------|------|--------|------|------------|
| | 1 | 2 | 3 | 4 | 5 |
| proefpersonen of stimuli | a | 0(2) | 1(30) | 0(-) | 0(-4) 0(L) |
| | b | 0(3) | 0(98) | 1(+) | 0(-3) 1(R) |
| | c | 1(4) | 0(97) | 1(+) | 1(3) 0(K) |
| | d | 1(4) | 1(101) | 0(-) | 0(-2) 1(R) |
| coderings- voorschrift | $\geq 4 = 1$ $\geq 100 = 1$ $+ = 1$ $\geq -1 = 1$ $R = 1$ $en \leq 35 = 1$ | | | | |

Bij deze coderingen behoeft geen rekening te worden gehouden met zogenaamde schaalnulpunten. In een Likert-item met de schaal 'zeer goed', 'goed', 'goed noch slecht', 'slecht', 'zeer slecht', kan de mediaan zowel liggen tussen 'zeer goed' en de rest; 'zeer goed', 'goed' en de rest; 'zeer goed', 'goed', 'goed noch slecht' en de rest; 'zeer goed', 'goed', 'goed noch slecht', 'slecht' en 'zeer slecht'. Verschuiving van het 'nulpunt' heeft geen invloed op de correlaties. Alle scores kunnen met een constante worden verhoogd of verlaagd.

Bij verzameling van de data in voldoende klassen is het ook mogelijk na te gaan of de relatie tussen de operator en de bedoelde klassificatie een niet-lineaire functie is. Door de operator dubbel te coderen kunnen acht verschillende relaties worden bepaald. De operator komt dan als twee variabelen in het onderzoek. De eerste variabele is een codering met 1 boven de mediaan en 0 beneden de mediaan. Bij de tweede wordt van de operator het hoogste en laagste kwartiel met 1 gecodeerd en de middelste twee kwartielen met 0. In tabel 4.3.2. staan de mogelijke relaties, die dan kunnen worden gevonden.

Tabel 4.3.2.

Mogelijke relaties bij dubbele codering van de operator

| | correlatie met bedoelde klassificatie | |
|-------------------------|---------------------------------------|------------|
| | mediaan | kwartielen |
| 1. positief lineair | $r > 0$ | $r = 0$ |
| 2. negatief lineair | $r < 0$ | $r = 0$ |
| 3. positief kromlijng | $r = 0$ | $r > 0$ |
| 4. negatief kromlijng | $r = 0$ | $r < 0$ |
| 5. positief accelerend | $r > 0$ | $r > 0$ |
| 6. negatief accelerend | $r > 0$ | $r < 0$ |
| 7. positief declinerend | $r < 0$ | $r > 0$ |
| 8. negatief declinerend | $r < 0$ | $r < 0$ |

Als waarschuwing moet nog worden gewezen op kunstmatige correlaties tussen de twee coderingen, wanneer de mediaancodering geen vijftig-vijftig procent is en/of de kwartielen codering niet vijftentwintig, vijftig, vijftentwintig procent, afwijkingen van vier procent of minder zijn bij niet te lage overeenkomstindices wel acceptabel. Het is ook mogelijk te werken met een pivot of opgegeven set waarbij alleen bijvoorbeeld de mediaancodering in is opgenomen.

Wanneer de gegevens in een klein aantal klassen zijn verzameld kan de eis van gelijke p-waarden moeilijkheden opleveren.

De volgende oplossingen zijn nu mogelijk.

1. Neem operatoren met extreem lage varianties niet in het onderzoek op. Operatoren met een p/q verdeling kleiner dan 10/90 procent geven veelal lage correlaties en kunnen daardoor weinig bijdragen aan het onderzoek. Variantie is een voorwaarde voor het berekenen van correlaties.
2. Geef een pivot op, waar operatoren met lage varianties niet in voorkomen
3. Ga na of een Q-analyse (zie paragraaf 5.2.) mogelijk is.

Theoretisch is na te gaan welke effecten optreden wanneer gewerkt wordt met ongelijke varianties. Wij herhalen daarvoor het bewijs uit paragraaf

3.4. dat $r_{is} = a_i$ echter niet met de aanname dat $\sigma_i = \sigma_j$. Uitgaande van ruwe scores is

$$r_{is} = \frac{\sum t_{bi} \cdot S_b - N \cdot p_i \cdot \bar{S}}{N \cdot \sigma_i \cdot \sigma_s}$$

Hierbij vindt de sommatie plaats over $b = 1$ t/m N (de elementen) en is

t_{bi} de uitslag van element b bij operator i ,

S_b de somscore over de set operatoren $i(1 \dots n)$,

\bar{S} de gemiddelde somscore,

p_i de proportie enen van operator i , en

σ_i en σ_s de standaarddeviaties van resp. i en s .

Nu is

$$S_b = \sum_{j=1}^n t_{bj}, \quad \text{dus} \quad t_{bi} \cdot S_b = t_{bi} \cdot \sum_{j=1}^n t_{bj}$$

verder is

$$\sum_{b=1}^N t_{bi} \cdot t_{bj} = N \cdot r_{ij} \cdot \sigma_i \cdot \sigma_j + N \cdot p_i \cdot p_j$$

dus

$$\sum_{b=1}^N t_{bi} \cdot S_b = N \cdot \sum_{j=1}^n r_{ij} \cdot \sigma_i \cdot \sigma_j + N \cdot p_i \cdot \sum_{j=1}^n p_j$$

Nu is

$$\sum_{j=1}^n p_j = \bar{S}$$

dus

$$r_{is} = \frac{\sum_{j=1}^n r_{ij} \cdot \sigma_i \cdot \sigma_j}{\sigma_i \cdot \sigma_j}$$

σ_i is een constante in teller en noemer en valt dus weg.

Verder is

$$\sigma_s = \sqrt{\sum_i \sum_j r_{ij} \cdot \sigma_i \cdot \sigma_j}$$

en dus

$$r_{is} = \frac{\sum_j r_{ij} \cdot \sigma_j}{\sqrt{\sum_i \sum_j r_{ij} \cdot \sigma_i \cdot \sigma_j}}$$

Nu is

$$r_{ij} = a_i \cdot a_j + a_{li} \cdot a_{lj} + \dots + a_{ki} \cdot a_{kj}$$

waarbij a_i de bedoelde lading is van operator i

a_{li} t/m a_{ki} de ladingen in de niet-bedoelde factoren die i gemeenschappelijk heeft met operator j .

$\sum_i \sum_j r_{ij} \sigma_i \sigma_j$ wordt dus

$$\sum_i \sum_j a_i \cdot a_j \cdot \sigma_i \cdot \sigma_j + \sum_i \sum_j a_{li} \cdot a_{lj} \cdot \sigma_i \cdot \sigma_j + \dots + \sum_i \sum_j a_{bi} \cdot a_{bj} \cdot \sigma_i \cdot \sigma_j$$

Nu is

$$\sum_i \sum_j a_i \cdot a_j \cdot \sigma_i \cdot \sigma_j \text{ gelijk aan } \left(\sum_j a_j \cdot \sigma_j \right)^2$$

Hetzelfde geldt voor de andere vormen, dus

$$\sum_i \sum_j r_{ij} \cdot \sigma_i \cdot \sigma_j = (\sum_j a_j \cdot \sigma_j)^2 + (\sum_j a_{1j} \cdot \sigma_j)^2 + \dots + (\sum_j a_{kj} \cdot \sigma_j)^2$$

r_{is} wordt dus gelijk aan

$$r_{is} = \frac{a_i \cdot \sum_j a_j \cdot \sigma_j + a_{1i} \cdot \sum_j a_{1j} \cdot \sigma_j + \dots + a_{ki} \cdot \sum_j a_{kj} \cdot \sigma_j}{\sqrt{(\sum_j a_j \cdot \sigma_j)^2 + (\sum_j a_{1j} \cdot \sigma_j)^2 + \dots + (\sum_j a_{kj} \cdot \sigma_j)^2}}$$

waarbij alle summaties plaats vinden over j (1 t/m n).

Bovenstaande formule is als volgt uit te schrijven:

$$r_{is} = a_i \cdot \left\{ \frac{1}{\sqrt{1 + (\sum_j a_{1j} \cdot \sigma_j)^2 / (\sum_j a_j \cdot \sigma_j)^2 + \dots + (\sum_j a_{kj} \cdot \sigma_j)^2 / (\sum_j a_j \cdot \sigma_j)^2}} \right. \\ + \frac{a_{1i} / a_i}{\sqrt{(\sum_j a_j \cdot \sigma_j)^2 / (\sum_j a_{1j} \cdot \sigma_j)^2 + 1 + \dots + (\sum_j a_{kj} \cdot \sigma_j)^2 / (\sum_j a_{1j} \cdot \sigma_j)^2}} \\ + \dots + \dots + \left. \frac{a_{ki} / a_i}{\sqrt{(\sum_j a_j \cdot \sigma_j)^2 / (\sum_j a_{kj} \cdot \sigma_j)^2 + (\sum_j a_{1j} \cdot \sigma_j)^2 / (\sum_j a_{kj} \cdot \sigma_j)^2 + \dots + 1}} \right\}$$

r_{is} is nu gelijk aan a_i als de vorm tussen haken gelijk is aan één.

Dit is het geval als

$$a_{1j} = a_{2j} = \dots = a_{kj} = 0 \quad (\text{voor alle } j, \text{ dus ook } i)$$

Verschillen in varianties tussen de operatoren is dus ernstig in die gevallen waar de set van operatoren i een intercorrelatiestructuur heeft van groepsfactoren en niet van belang wanneer deze structuur bestaat uit één algemene factor zonder groepsfactoren. Door nu clusters te vormen

waar bij ieder geval altijd sprake is van één algemene factor kan gedeeltelijk bovenstaande problematiek worden ondervangen. Dit laatste vindt nu bij toepassing van de clusteranalyse plaats. Bij deze analyse wordt uitgegaan van één variabele en dit cluster uitgebreid met die variabele die het hoogst met de eerste correleert, onder voorwaarde dat deze correlatie ligt boven de overeenkomstindex. Aan dit cluster van twee variabelen wordt weer een derde variabele toegevoegd en wel diegene die het hoogst met de somscore van de eerste twee correleert, onder dezelfde voorwaarde. Deze bewerking wordt net zolang uitgevoerd tot er geen variabelen meer zijn, die voldoen aan de gestelde voorwaarde. Hierdoor is het niet mogelijk een cluster van variabelen te krijgen waarin geen algemene factor aanwezig is.

Uit bovenstaande formule is verder af te lezen dat schatting van a_i met behulp van r_{is} bij grote verschillen in de varianties van de operatoren vooral moeilijkheden geeft bij operatoren die een positieve lading hebben in de algemene factor en negatieve ladingen in de niet-bedoelde factoren. In dat geval is er sprake van een duidelijke onderschatting van a_i bij berekening van r_{is} . Dergelijke operatoren hebben een belangrijke 'surpressor' functie en kunnen moeilijk worden gemist. Een oplossing zou kunnen worden gevonden door niet met te hoge overeenkomstindexen te werken. Deze conclusie is echter tegengesteld met de eerste. Bij hoge overeenkomstindices is er een veel grotere zekerheid, dat de cluster een sterk dominerende algemene factor heeft en dus tegengestelde ladingen een veel minder belangrijke rol spelen.

Onze conclusie is dan ook, dat het niet voldoen aan de voorwaarde van gelijke varianties vooral ernstige fouten in de schattingen van de ladingen geeft bij ingewikkelde structuren van intercorrelaties tussen de operatoren; veel factoren die duidelijk overlappen met positieve en negatieve ladingen.

De problematiek is tot een bepaalde hoogte te voorkomen door niet met te lage overeenkomstindices te werken. Omgekeerd behoeft de voorwaarde niet zo ernstig genomen te worden in die gevallen waar sprake is van eenvoudige structuren van weinig factoren die nagenoeg niet overlappen.

4. De overeenkomstindex

In paragraaf 3.5. en wel in aansluiting op paragraaf 2.5. 'De orde van de klassificatie' is gesproken over de overeenkomstindex. Door het vaststellen van deze index worden de clusters gedefinieerd. Meestal wordt gewerkt met de index of grens van 0,30. Dit betekent dat een operator, waarvan de klassificatie 0,30 of meer correleert met de bedoelde klassificatie, voldoende valide is. Bij ieder onderzoek blijken er meer of minder optimale waarden voor de overeenkomstindex te zijn. Zo bleek Hermans (1967) voor de constructie van een prestatie-motivatieschaal beter te kunnen werken met 0,20, terwijl Pennings (1966) bij 0,30 goede resultaten kreeg. Daar tegenover gaf de analyse van opstellen een pregnanter resultaat bij een grens van 0,40 (Kempen en Boon van Ostade, 1968, zie ook paragraaf 5.2.4.), terwijl een analyse van beoordelingen van de noodzakelijke eigenschappen van de ondergrondse mijnarbeider door van Hecke (1966) pas bij een grens van 0,70 aparte clusters gaf.

De indruk is, dat de validiteit van de gemiddelde operator voor de bedoelde klassificatie bepalend is voor de grootte van de optimale grens. Deze is laag bij vragenlijstitems voor een vrij abstract construct als prestatie-motivatie en hoog voor oordelen over de noodzakelijke eigenschappen van een functie bij beoordelaars die de functie zeer goed kennen (van Hecke, 1966; in dit geval opzichters van één bedrijf die zelf jarenlang als mijnwerker werkzaam waren geweest).

Getracht is te werken met een glijdend criterium, dat wil zeggen een criterium dat zich aanpast aan de resultaten zoals deze tijdens de analyse blijken. In dat geval is het niet nodig om de analyse te herhalen met een lager c.q. hoger criterium als blijkt dat het

gestelde niet voldoet. Hiervoor is in de eerste plaats het varimax-criterium gekozen. Dit criterium is een objectieve formulering van Thurstone's 'simple structure' principe door Kaiser (1958) en komt hierop neer dat getracht wordt de variantie van de gekwadrateerde ladingen per factor te maximaliseren, waardoor er zoveel mogelijk hoge en lage ladingen per factor komen. Dit criterium bleek niet te voldoen. Er kwamen teveel kleine clusters. Eén van de redenen hiervan is gelegen in het feit dat niet gewerkt kan worden met de 'normal varimax' formule, omdat de communaliteiten (h^2) niet bekend zijn. Wherry en Winer (1953) hebben wel een methode ontwikkeld om de h^2 te schatten. Deze methode voldoet echter niet bij een door ons uitgevoerde toepassing (Boon van Ostade, 1963) en wel omdat de h^2 geschat wordt per factor in plaats van over alle factoren tegelijk. Betere methoden voor h^2 schattingen o.a. de gekwadrateerde multiple correlatie (Wrigley, 1957) zijn niet toegepast, omdat deze methoden moeilijk toepasbaar zijn bij veel variabelen, juist één van de sterke punten van de iteratieve clusteranalyse.

Ook is gewerkt met de Kuder-Richardson formule 20 en wel in navolging van Elshout's hiërarchische clusteranalyse (1963). Ook dit criterium voldoet niet en wel om twee redenen. Met dit criterium bestaat het kleinste cluster altijd uit twee operatoren. Bij een specifiek cluster van één operator is de K.R.(20) niet te berekenen. Verder geeft dit criterium ook veel kleine clusters. Ter illustratie staan in tabel 4.4.1. de uitkomsten van deze formule bij de verschillende stadia van het eerste A.M. cluster (zie paragraaf 4.5.).

Tabel 4.4.1.

K.R.(20) waarden van het eerste A.M. cluster
in de verschillende stadia van zijn opbouw.

| stadia | cluster operatoren | K.R.(20) |
|--------|---------------------------|----------|
| 2 | 5 + 1 | 0,81452 |
| 3 | 5 + 1 + 2 | 0,81450 |
| 4 | 5 + 1 + 2 + 4 | 0,839 |
| 5 | 5 + 1 + 2 + 4 + 7 | 0,847 |
| 6 | 5 + 1 + 2 + 4 + 7 + 8 | 0,852 |
| 7 | 5 + 1 + 2 + 4 + 7 + 8 + 9 | 0,836 |

Bij gebruik van de K.R. formule als overeenkomstindex was de analyse gestopt bij het cluster 5 + 1.

Het niet voldoen van bovengenoemde twee criteria is in het licht van de in het vorige hoofdstuk ontwikkelde theorie niet zo verwonderlijk. De clusteranalyse streeft niet naar een uni-factor model, hetgeen het varimax criterium wel doet (zie paragraaf 3.5.) en de definitie van de meetfout is, zoals die bij de toepassing van de K.R.(20) formule plaats vindt, niet in overeenstemming met die van de clusteranalyse (zie paragraaf 3.6.).

Bij duidelijke onzekerheden ten aanzien van de overeenkomstindex, blijkt praktisch de beste strategie om te beginnen met lage waarden (0,20 - 0,30) en dan afhankelijk van de mate waarin de ladingen van de operatoren in één cluster onderling verschillen deze grens te verhogen. Ook een goede aanwijzing vormen de correlaties van de operatoren vóór hun opname. Wanneer de eerste operatoren correlaties hebben in de orde van grootte van 0,50 of hoger en de latere van 0,30 of lager, is er een duidelijke aanwijzing, dat een strengere grens betere resultaten geeft. Optimale presentaties zijn die, waar de ladingen van de cluster operatoren homogeen zijn (zie regel d. paragraaf 3.4.).

Een ondergrens van de overeenkomstindex wordt gegeven door de grootte van de steekproef. Wanneer de klassificaties van een operator en de bedoelde klassificatie zo weinig overeenkomen, dat de nulhypothese niet kan worden verworpen, kan niet over een valide operator worden gesproken. Dit betekent dat de overeenkomstindex bij steekproeven groter of gelijk aan 30 altijd groter of gelijk moet zijn aan (zie bijlage I.1., geval 1a en 1b)

$$r_c \geq \frac{1,96}{\sqrt{N}}$$

of bij steekproeven kleiner dan 30

$$r_c \geq \frac{t'}{\sqrt{(t')^2 + N - 2}}$$

waarbij r_c de overeenkomstindex is,

N de steekproefgrootte, en

t' de student's t bij $N - 2$ vrijheidsgraden

Het gebruik van bovenstaande minimale overeenkomstindices sluit niet uit, dat in de analyse correlaties en dus clusters worden gevonden, die door toeval ontstaan. Dit is het gevolg van de herhaling van de bewerking. Bovengenoemde toetsingen gaan uit van een éénmalige berekening. In paragraaf 4.6. wordt verder op deze problematiek ingegaan.

Concluderend kan worden gezegd, dat de onderzoeker de keuze van de overeenkomstindex kan oplossen door de analyse een aantal malen uit te voeren met indices, die steeds strenger worden, en wel afhankelijk van de verkregen resultaten.

4.5. Vorming van clusters

Voor het vormen van een cluster, dat beantwoordt aan de voorschriften, zoals uiteengezet in paragraaf 3.4., gaat de analyse als volgt te werk:

- Uitgaande van één operator wordt nagegaan hoe de overeenkomsten zijn van de klassificatie van deze operator met die van de andere operatoren.
- De operator, die de grootste overeenkomst heeft met de eerste operator vormt nu samen met deze een cluster onder voorwaarde dat de overeenkomst ligt boven de gestelde overeenkomstindex.
- De ongewogen somscore van deze eerste twee operatoren wordt weer vergeleken met de klassificaties van alle andere operatoren.
- De operator, die nu de grootste overeenkomst heeft met de gevormde somscore, wordt bij dit cluster opgenomen, weer onder voorwaarde dat de overeenkomst ligt boven de overeenkomstindex.
- Bovengenoemde bewerking wordt net zo lang uitgevoerd tot dat er geen operatoren zijn om op te nemen of geen één operator boven de gestelde index overeenkomt met de somscore.

De overeenkomst tussen de operatoren en de ongewogen somscore wordt met behulp van produktmoment correlaties bepaald als maat voor de validiteit van de operator (zie paragraaf 3.1.). De bedoelde klassificatie wordt dus uitgaande van één operator opgebouwd met toepassing van de regel dat de correlatie groter of gelijk moet zijn aan de minimumwaarde, zoals gesteld door de onderzoeker de z.g. overeenkomst-index. Hierdoor wordt het mogelijk te voldoen aan de voorwaarde, dat de eerste centroïd van de intercorrelaties van de opgenomen operatoren een algemene factor is (de convergerende validiteit, zie paragraaf 3.5.).

Door berekening van de correlaties van de op deze wijze gevormde clusterscore met de operatoren of clusters van niet-relevante condities kan worden nagegaan of deze algemene factor ook de bedoelde factor is (de divergerende validiteit, zie paragraaf 3.5.).

De bewerking wordt nu *controleerend* uitgevoerd, als bovengenoemde clustervorming alleen plaats vindt voor de operatoren van de gestelde pivot of het apriori cluster, die door de onderzoeker voor de analyse is bepaald. Operatoren, die niet in de pivot voorkomen, kunnen dus dan niet worden opgenomen in het cluster. De bewerking wordt *toetsend* uitgevoerd als bij de vorming van ieder volgend cluster na het eerste de operatoren van voorafgaande clusters wel opgenomen kunnen worden. Bij *controleerend-exploratief* onderzoek is dit niet mogelijk, evenmin als bij *explorerend* onderzoek. De analyse vindt *explorerend* plaats, wanneer er geen pivots zijn opgegeven. In dat geval worden alle operatoren of variabelen als één pivot opgevat.

Bovenstaande bewerking verschilt van andere (onder andere Elshout's hiërarchische clusteranalyse, 1963), doordat gewerkt wordt met een ongewogen somscore en niet bijvoorbeeld met de gemiddelde correlatie van de clusteroperatoren en een eventueel op te nemen operator. De reden van dit verschil is gelegen in de definiëring van de bedoelde klassificatiemaat, als ongewogen som van operatoren met gelijke varianties (zie paragraaf 3.3.) en dat deze somscore een schatting is van de factorscore van de eerste centroïd (zie paragraaf 3.4.). De operatoren die opgenomen zijn in het cluster, voldoen aan de eis van de convergerende validiteit en kunnen dus voorlopig afhankelijk van een nog nader uit te voeren controle op de divergerende validiteit, beschouwd worden als adequate operatoren voor de bedoelde klassificatie.

Een belangrijke voorwaarde voor de toepassing van het centroid model is dat de som van de intercorrelaties van de operatoren, die het cluster vormen, niet nul is (zie paragraaf 3.4. voorwaarde 4). Dit kan plaats vinden wanneer met te lage overeenkomstindices wordt gewerkt. Wanneer echter als minimumgrens van deze indices de voor 5% significante correlatie wordt genomen (zie paragraaf 4.4.) is de kans, dat niet aan deze voorwaarde wordt voldaan, verwaarloosbaar klein en zeker uitgesloten wanneer de gevonden clusters worden gecontroleerd op hun kans als toevalsresultaat (zie paragraaf 4.6.). Ernstiger wordt deze kwestie, wanneer operatoren met negatieve en positieve intercorrelaties het cluster vormen. Dit moet worden voorkomen. Daarom worden operatoren met negatieve correlaties pas gebruikt voor de vorming van de clusterscore nadat ze zijn *gespiegeld*. In paragraaf 3.1. is op deze werkingmogelijkheid gewezen. Het coderingsvoorschrift van deze operatoren wordt omgedraaid; de nul-coderingen worden één-coderingen en omgekeerd. De verantwoording van deze spiegeling berust op het feit dat dergelijke operatoren wel degelijk informatief zijn voor de bedoelde klassificatie, ondanks het feit dat de kans op fouten duidelijk groter is dan de kans op treffers, en wel in verschil met operatoren waar beide kansen nagenoeg gelijk zijn.

Er wordt nog op gewezen, dat bovengenoemde bewerking (de beslissing of de operator tot het cluster behoort) berust op niet-'spurious' zijnde correlaties. Dit is een belangrijk verschil met de iteratieve factoranalyse van Wherry en Gaylord (1943). Bij deze analyse wordt deze beslissing in eerste instantie genomen met behulp van de correlaties tussen de operatoren en de somscore van de gehele

pivot. Dergelijke correlaties zijn overschattingen van de werkelijke correlaties en wel des te meer naarmate de pivot kleiner is, omdat de operatoren zelf de somscore bepalen en dus gedeeltelijk met zichzelf worden gecorreleerd. Roskam heeft een verbeterde versie van deze analyse ontwikkeld. De correlaties worden in die analyse berekend tussen de operator en alle andere van de pivot en zijn dan niet meer 'spurious'. Het nadeel van deze methode is echter, dat de operatoren dan steeds met een andere somscore worden vergeleken, waardoor de clustervorming een proces kan worden zonder eind. In de eerste bewerking worden dan bijvoorbeeld bepaalde operatoren als onvoldoende beschouwd, die dan in verdere bewerkingen weer voldoende blijken te zijn wat het gevolg heeft dat andere operatoren weer onvoldoende worden, waardoor weer de eerstgenoemde operatoren onvoldoende blijken te zijn enz. Een eeuwig doorlopend, niet-convergerend proces van clustervorming vindt dan plaats. Vooral bij het gebruik van computers moet een dergelijke bewerking worden voorkomen.

Door steeds die operator te kiezen, die de hoogste correlatie met de clusterscore heeft, wordt ook zorg gedragen, dat regel b van paragraaf 3.4. waarin gesteld wordt dat gestreefd moet worden naar een zo groot mogelijke som van intercorrelaties van de operatoren, is gewaarborgd. Deze regel dient om de schattingsfout van de factorscore van de eerste centroïd door middel van de clusterscore zo klein mogelijk te houden (zie formule 12). Door deze regel te volgen wordt ook de psychometrische betrouwbaarheid van het cluster verhoogd (zie paragraaf 3.6.). In paragraaf 3.2. is uiteengezet, dat het primair stellen van deze regel en dus de psychometrische betrouwbaarheid, zoals gebruikelijk bij de methoden van itemanalyses, ten koste kan gaan van de representativiteit. Om deze reden is in de bewerking van de iteratieve cluster-

analyse deze regel niet centraal gesteld. Bij methoden die niet met een somscore, maar met de intercorrelaties tussen de operatoren rekening houden, is dit meestal wel het geval.

Een probleem vormt nog met welke operator moet worden begonnen. Gebleken is, dat hiervoor het beste die operator gekozen kan worden, die het hoogst met de somscore van alle operatoren van de pivot correleert. Onderzocht is of deze correlaties uitgerekend moesten worden tussen de operator en alle andere of de operator en alle operatoren (de operator zelf inbegrepen). 'Part-whole' correlaties zijn 'spurious' hoog en geven minder goede informaties. Gebleken is echter dat de tweede methode het beste voldoet (berekening van de correlaties zonder correctie voor 'spuriousness'). De reden hiervan is gelegen in het feit dat bij uitsluiting van de operator zelf vooral die operatoren worden gekozen die positieve en negatieve ladingen hebben in verschillende factoren. De correlaties van deze operatoren zijn echter slechte schattingen van de ladingen als er verschillen tussen de varianties van de operatoren zijn (zie paragraaf 4.3.).

Een verbetering van de keuze van de uitgangsvariabele is gevonden door toepassing van een spiegeling, zoals ook door Thurstone (1947, p.153) bij de centroïd methode wordt uitgevoerd. Eerst wordt nagegaan of er operatoren zijn die negatief correleren met de somscore, vervolgens welke operator het hoogst negatief correleert. Deze operator wordt dan gespiegeld: zijn 1-waarden worden 0, zijn 0-waarden 1. De nieuwe somscore wordt berekend en weer de correlaties van de operatoren met deze somscore. De operator met de hoogste negatieve correlatie wordt weer gespiegeld enz. Deze bewerking stopt als er geen negatieve corre-

laties zijn of als er meer dan driekwart van de operatoren zijn gespiegeld.

De beschreven clustering zullen we nader verduidelijken met een voorbeeld. Het ging hierbij om de mate van overeenkomst te analyseren tussen klassificaties van intelligentietests, beoordelingen en niveau's van de vooropleiding bij 49 aspirant mijnscholieren (A.M.-ers). In tabel 4.5.1. staan de operatoren van dit onderzoek.

Tabel 4.5.1.

Operatoren A.M. onderzoek

1. Beoordeling 'algemene intelligentie' door psycholoog en assistent op grond van alle gegevens.
2. Beoordeling 'algemene intelligentie' door psycholoog op grond van alleen een gesprek.
3. Beoordeling 'algemene intelligentie' door assistent op grond van alleen een gesprek en observaties in discussiegroepen en buitenproeven.
4. Verstandelijke Leeftijd Ontwikkeling Test (VLOT).
5. Algemene Intelligentie test van Coetsier (C.I.).
6. Progressive Matrices, 1938.
7. Letterreeksen A
8. Letterreeksen a
9. Niveau van de vooropleiding.

In tabel 4.5.2. staan de nul en één coderingen van de 49 aspirant mijnscholieren.

Tabel 4.5.2.

Nul en één matrix A.M. onderzoek

| Operator | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Somscore |
|--------------|----|---|---|---|---|---|---|---|---|---|----------|
| Proefpersoon | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 9 |
| | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| | 3 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 5 |
| | 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 9 |
| | 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 8 |
| | 6 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 8 |
| | 7 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 7 |
| | 8 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 2 |
| | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| | 10 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| | 11 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| | 14 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 5 |
| | 15 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 6 |
| | 16 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 4 |
| | 17 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 8 |
| | 18 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 2 |
| | 19 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 3 |
| | 20 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 2 |
| | 21 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 6 |
| | 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 23 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 8 |
| | 24 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 6 |
| | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| | 26 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 27 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 |
| | 28 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 3 |
| | 29 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 2 |
| | 30 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 6 |
| | 31 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 7 |
| | 32 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 33 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 4 |
| | 34 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 5 |
| | 35 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| | 36 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 5 |
| | 37 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| | 38 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 9 |
| | 39 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 4 |
| | 40 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 7 |
| | 41 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 6 |
| | 42 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 5 |
| | 43 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 4 |
| | 44 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 3 |
| | 45 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 5 |
| | 46 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 8 |
| | 47 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 3 |
| | 48 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 8 |
| | 49 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 6 |

De somscore voor proefpersoon 1 bij alle operatoren is 9, voor proefpersoon 2: 1, proefpersoon 3: 5. Er staan resp. 9, 1 en 5 één coderingen in de betreffende rijen. Berekening van de point-biserial r tussen de somkolom en de operator kolom 1 t/m 9 geeft de volgende negen correlaties: 1; 0,74, 2; 0,74, 3; 0,43, 4; 0,74, 5; 0,82, 6; 0,37, 7; 0,69, 8; 0,61, 9; 0,49. De uitgangsvariabele is dus 5; de Algemene Intelligentietest van Coetsier. In tabel 4.5.3. staan in de eerste kolom de correlaties van operator 5 met de andere operatoren. De hoogste correlatie is die van 1; de beoordeling van de psycholoog en assistent. De tweede kolom geeft de correlaties van 5 en 1 met de andere operatoren. Operator 2 geeft de hoogste correlatie; de beoordeling van de psycholoog. De derde kolom geeft de correlaties van 5, 1 en 2. De hoogst correlerende is nu 4; de VLOT. De vierde kolom wijst 7 aan; de Letterreeksen A. De vijfde 8; de Letterreeksen a en de zesde 9; het niveau van de vooropleiding. Deze correlatie is echter vrij laag. Voor testcores en beoordelingen is een overeenkomstindex van 0,40 tot 0,60 het meest aangewezen. Dit blijkt ook bij opname van operator 9. De ladingen van het cluster in de zevende kolom lopen dan van 0,86 (74%) tot en met 0,50 (25%).

Tabel 4.5.3.

De correlaties bij de verschillende stadia van het eerste A.M. cluster

| stadia | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|----------|---|--------|--------|--------|--------|--------|--------|--------|
| operator | 1 | 0,68 | (0,91) | (0,82) | (0,79) | (0,79) | (0,79) | (0,77) |
| | 2 | 0,67 | 0,60 | (0,82) | (0,82) | (0,81) | (0,68) | (0,61) |
| | 3 | 0,00 | 0,00 | 0,23 | 0,28 | 0,30 | 0,29 | 0,30 |
| | 4 | 0,51 | 0,55 | 0,62 | (0,79) | (0,74) | (0,72) | (0,72) |
| | 5 | (1,00) | (0,92) | (0,92) | (0,88) | (0,87) | (0,87) | (0,86) |
| | 6 | 0,00 | 0,29 | 0,18 | 0,25 | 0,17 | 0,15 | 0,15 |
| | 7 | 0,56 | 0,58 | 0,60 | 0,57 | (0,72) | (0,72) | (0,72) |
| | 8 | 0,52 | 0,55 | 0,52 | 0,51 | 0,53 | (0,68) | (0,63) |
| | 9 | 0,18 | 0,30 | 0,33 | 0,36 | 0,36 | 0,33 | (0,50) |

Duidelijk is aan de correlaties af te lezen dat de clustering de algemene variantie versterkt (operator 4 gaat van 0,51 naar 0,62, operator 9 van 0,18 naar 0,36) ten koste van de specifieke (operator 5 gaat van 1,00 naar 0,86, operator 1 van 0,91 naar 0,77).

Het resultaat van de analyse is een bevestiging van de bevindingen van Meehl (1954) ten aanzien van klinische predicties. De begripsvaliditeit van de psychologenbeoordeling (operator 2) is lager of gelijk aan die van de tests. Een bevestiging van de juistheid van deze interpretatie wordt ook verkregen door de correlatie met het niveau van de vooropleiding (operator 9). Het is echter mogelijk dat deze correlatie is ontstaan door contaminatie met operator 1. Bij opname van deze operator in het cluster (stadium 2) gaat de correlatie van 9 van 0,18 naar 0,30. In de volgende stadia wordt echter deze correlatie niet duidelijk meer verhoogd. Aan deze predictieve validiteit kan dus niet veel waarde worden gehecht. Ook de beoordeling op grond van alle gegevens (operator 1) heeft niet de hoogste lading. De intelligentietest van Coetsier voldoet in dat opzicht beter. Dit kan komen doordat de assistenten duidelijk een andere opvatting hebben over 'algemene intelligentie' (operator 3). Opvallend zijn namelijk de lage correlaties van de beoordelingen van de assistent (3) en van de Progressive Matrices (6). Het is mogelijk dat deze een aparte cluster vormen. Door de analyse te herhalen maar nu alleen op de overgebleven operatoren (3, 6 en 9), kan worden nagegaan of er inderdaad sprake is van een tweede A.M.-cluster. Hierbij dienen echter wel de operatoren van het eerste cluster te worden uitgesloten, anders is de kans groot dat het eerste cluster weer terug komt. Door berekening van de correlaties van de uitgesloten operatoren kan worden nagegaan in hoeverre zij ladingen hebben in het tweede cluster, zgn. tussenvariabelen zijn.

De mate van uitsluiting van operatoren die al in clusters zijn opgenomen, is, zoals al aangegeven, verschillend voor exploratief- en toetsingsonderzoek. In het eerste geval kan geen enkele operator, die al opgenomen is, in een ander volgend cluster komen. Hierdoor wordt een duidelijk inzicht in de structuur van de intercorrelaties verkregen. De clusters zijn alle in voldoende mate onafhankelijk. Bij toetsing gaat het meestal niet om exploratie van de structuur, maar om het beste cluster. Om die reden wordt in deze vorm van de analyse alleen voor de keuze van de uitgangsvariabele operatoren, die al zijn opgenomen, uitgesloten. Meestal is dan het eerste cluster het bedoelde. Soms echter is een tweede cluster toch beter: bijvoorbeeld doordat er meer operatoren in zijn opgenomen en daardoor het cluster representatiever is. Soms ook heeft een volgend cluster een hogere K.R.(20). De kans dat het laatstgenoemde resultaat wordt verkregen is echter bijzonder klein: 1 tot 2 maal op de twintig toetsingen. Blijkbaar heeft de spiegelmethode voor de keuze van de uitgangsvariabele een grote mate van trefzekerheid.

4.6. Controle van de clusters

De evaluering van de verkregen clusters kan allereerst geschieden op grond van het aantal operatoren dat in ieder cluster is opgenomen. Het grootste cluster is meestal het belangrijkste, het bedoelde. Uiteraard is dit niet altijd het geval. Controle moet plaats vinden op grond van de inhoud van de operatoren. Het kan zijn dat de operatoren een niet-bedoeld aspect gemeenschappelijk hebben in tegenstelling met operatoren, die niet in het cluster zijn opgenomen. Dit zijn dan meestal methodische clusters als bijv. alle negatief gestelde vragen, alle moeilijke vragen, alle 'open-end' vragen of alle aan het eind van het onderzoek gestelde vragen.

Voortdurend moet men bedacht zijn op dergelijke clusters. Ter adstructie worden hier drie onderzoekingen gereleveerd, waar een methodische cluster voorkwam. Bij de z.g. kapittelenquête (zie paragraaf 4.1.) viel de N.I.P.G. Satisfactieschaal in eerste instantie uiteen in twee clusters van alle positief gestelde- en van alle negatief gestelde vragen (Molenaars en Jansen van der Slichte, 1968). Dit kwam doordat de niet-beantwoorde vragen steeds dezelfde codering hadden gekregen, in plaats van bijvoorbeeld een aselechte codering van nul en één. Een onderzoek bij 100 bakkers gaf als eerste en grootste cluster alle 'open-end' vragen van de enquête, gecodeerd met 'antwoord gegeven' (1) en 'geen antwoord gegeven' (0). Hierbij is duidelijk sprake van een instelling tegenover het onderzoek. Er zijn personen die gesloten zijn en weinig zeggen, terwijl anderen altijd een antwoord geven. Starren (1965) kreeg bij 105 jongens uit de eerste klas van de Technische Vakschool (Chemische Bedrijven, N.V. Ned. Staatsmijnen) een tweede cluster van 24 vragen, waar geen enkel criterium

(cijfers, tests en beoordelingen) mee correleerde. Bij nadere inspectie van het cluster viel op, dat er veel vragen (17 van de 24) voorkwamen uit de autobiografische vragenlijst, die in het onderzoek, dat gemiddeld twee uur duurde, aan het eind was afgenomen. De vragen betroffen verder topics, die al in de voorafgaande twee vragenlijsten aan bod waren geweest. De conclusie was dat dit cluster denkelijk het gevolg was van een stereotype wijze van beantwoording tengevolge van een toegenomen weerzin voor de vragen.

Een tweede formele controle geeft de uitkomst van de K.R.(20) of (21) formule. Deze moet minstens significant zijn (zie bijlage I.4.). Rekening moet daarbij gehouden worden met het feit, dat deze formules te hoge waarden geven (zie paragraaf 3.6.). Als significantiegrens dient dus minstens de 1% drempel te worden gebruikt (zie ook paragraaf 4.1.). Absoluut gezien zijn clusters met K.R.(20) waarden beneden 0,50 vrij dubieus en moet er aan dergelijke clusters niet veel waarde worden gehecht.

Naast de K.R.(20) formule is vooral bij controle-onderzoek de divergerende validiteit belangrijk. De correlaties tussen de gevonden clusters moeten voldoende laag zijn in ieder geval lager dan de gebruikte overeenkomstindex.

Daarnaast moet gekeken worden of de ladingen van operatoren niet te veel variëren (zie bijlage I.2.). Als dit het geval is, is er meestal met een te lage overeenkomstindex gewerkt. Bij een te hoge overeenkomstindex daarentegen komen er teveel specifieke en kleine clusters. Een specifiek cluster bestaat maar uit één operator; er correleert geen één andere operator met de betreffende boven de gestelde overeenkomstindex, of alleen operatoren die al in een cluster zijn opgenomen. Het is verstandig aan dergelijke clusters geen waarde te hechten

Bij exploratief onderzoek kan het voorkomen, dat kleine clusters onvoldoende tot hun recht komen. Operatoren namelijk, die in voorafgaande clusters voorkomen, kunnen niet meer in volgende clusters worden opgenomen. Deze werkwijze is noodzakelijk om onafhankelijke clusters te verkrijgen. Inspectie van de correlaties van alle operatoren (ook de niet opgenomen) leert welke operatoren tussenvariabelen zijn d.w.z. ladingen hebben in meer dan één cluster. Door nu deze operatoren samen met die van het cluster in een tweede bewerking als pivot op te geven, kunnen dergelijke clusters beter worden gepresenteerd. Een correlatie met een cluster betekent nog niet dat de operator tot het cluster kan worden gerekend. Niet-voorziene negatieve correlaties met andere niet opgenomen, maar wel en hoger correlerende operatoren, kunnen tot gevolg hebben dat de operator toch niet voldoet.

Vooraf wanneer het resultaat bestaat uit veel kleine clusters dient men bedacht te zijn op een toevalsresultaat. Een dergelijk resultaat kan worden verkregen bij aselechte data. Vooraf bij onderzoeken met veel meer operatoren dan elementen is het aan te bevelen de data hierop te controleren. Bij een serie beslissingen, zoals in de iteratieve clusteranalyse plaats vindt, kunnen namelijk - ook als wordt gewerkt met een 1% significantiedrempel voor de overeenkomst-index - resultaten worden verkregen, die specifiek zijn voor de onderzochte steekproef en die niet kunnen worden gegeneraliseerd.

Het aantal beslissingen dat in de analyse plaats vindt, is gelijk aan $\frac{1}{2}n(n-1)$, waarbij n het aantal operatoren is. Iedere operator wordt vergeleken met alle andere operatoren, waarvan deze vergelijking nog niet heeft plaats gevonden. Het aantal toetsingen is dus de som van de reeks: $(n-1)$, $(n-2)$, ... $(n-n-1)$, $(n-n)$. Deze som is

gelijk aan $\frac{1}{2}n(n-1)$. Door herhaalt te toetsen (te vergelijken met de overeenkomstindex) kunnen 'significante' correlaties voorkomen, die geen betekenis hebben. Het is nu mogelijk om dit aantal te berekenen met behulp van het binomiale model, zoals beschreven door Jones en Fiske (1953). Met dit model kan achteraf - nadat de analyse is uitgevoerd - na worden gegaan in hoeverre het totale resultaat van de analyse toevallig genoemd moet worden. Het is ook mogelijk met het binomiale model bepaalde gedeelten van de resultaten te toetsen op hun toevalligheid. Bij kleine aantallen toetsingen is het echter adequater om het chi-kwadraat model toe te passen, omdat dit model scherper is: het houdt rekening met de grootte van de gevonden correlaties en dichotomiseert ze niet in 'significant' (groter dan de overeenkomstindex) en 'niet significant' (kleiner dan de overeenkomstindex). De chi-kwadraat toets is echter veel bewerkelijker, terwijl bovengenoemde bezwaar bij veel operatoren (meer dan 20) niet zo ernstig is.

Een probleem wordt gevormd door de eis van onafhankelijkheid. De toets gaat ervan uit, dat voor iedere toetsing een nieuwe steekproef is getrokken. Dit is bij de clusteranalyse niet het geval. Er wordt steeds met één en dezelfde steekproef gewerkt. Jones en Fiske (1953, p.378 en 9) achten dit probleem niet zo ernstig voor een intercorrelatiematrix, waar $\frac{1}{2}n(n-1)$ intercorrelaties van n variabelen worden bekeken. Hun conclusie is: "the binomial test, if interpreted with caution, may serve as an approximate index of independence among the n variables" (p.379). Zij hebben deze conclusie in drie onderzoeken empirisch bevestigd gevonden. Hetzelfde resultaat is ook hier verkregen bij de simulatie-onderzoeken (zie paragraaf 6.1.). Bij de iteratieve clusteranalyse wordt weliswaar niet uitgegaan van de intercorrelatie-

matrix, echter ook $\frac{1}{2}n(n-1)$ maal getoetst. Bovenstaande conclusie kan dus ook voor deze analyse worden gesteld. De juistheid hiervan is gecontroleerd in twee simulatie-onderzoeken (zie paragraaf 6.2.2.) en bevestigd bevonden.

Uitvoering van de binomiale toets is bij de clusteranalyse echter niet uitvoerbaar, als het aantal toetsingen vermenigvuldigt met de significantiegrens niet duidelijk minder is dan het aantal operatoren. Het is namelijk in deze analyse niet mogelijk om meer dan $(n-1)$ 'significante' correlaties te vinden. Hierdoor wordt afhankelijk van het aantal operatoren de minimale significantiegrens vastgelegd. Deze significantiegrens bepaalt nu weer samen met de grootte van de steekproef de minimale overeenkomstindex.

In tabel 4.6.1. is dit nader uitgewerkt en wel met een benaderingsformule waarbij

$$(13) \quad N \geq \left\{ \frac{t\left\{\alpha \geq \frac{1}{n-1}\right\}}{z(r_c)} \right\}^2 + 3$$

Hierbij is N de grootte van de steekproef,

t de standaardnormale t -waarde bij α ,

α de significantiegrens,

n het aantal operatoren,

z de Fisher z -waarde, en

r_c de overeenkomstindex

Tabel 4.6.1.

Minimale overeenkomstindex voor toetsing van de resultaten van de clusteranalyse op toeval bij verschillende steekproefgrootten en verschillende aantallen operatoren.

| aantal operatoren | grootte van de steekproef | | | | | |
|------------------------|---------------------------|---------|--------|-------|-------|-------|
| 2-10 | 3-71 | 3-43 | 3-31 | 3-18 | 3-12 | 3-9 |
| 11-20 | 72-99 | 44-60 | 32-43 | 19-25 | 13-16 | 10-11 |
| 21-100 | 100-169 | 61-101 | 44-72 | 26-41 | 17-25 | 12-17 |
| 101-200 | 170-199 | 102-119 | 73-86 | 42-48 | 26-29 | 18-20 |
| 201-1000 | 200-272 | 120-164 | 87-115 | 49-65 | 30-39 | 21-26 |
| overeenkomst- index | 0,20 | 0,25 | 0,30 | 0,40 | 0,50 | 0,60 |

De gebruikte formule voor tabel 4.6.1. die uitgaat van Fisher's z-coëfficiënt is gecontroleerd met de exactere formule

$$N \geq t_1^2 \{ \alpha \geq 1/n-1; df = N-2 \} (1/r_c^2 - 1) + 2$$

waarbij t_1 de student's t is met als vrijheidsgraden (df) N-2.

Hierbij bleek dat de uitkomsten van deze formule nagenoeg overeen kwamen met de benaderingsformule zodat er geen bezwaar is om formule 13 te gebruiken, die geen iteraties behoeft.

Als voorbeeld zal bovengenoemde toetsing hier worden uitgevoerd voor de resultaten van de tiende aselekt getrokken groep zusters van de z.g. kapittel-enquête (zie paragraaf 4.1.). Bij een geheel toevallige resultaat zouden bij 15 vragen, volgens de formule voor het gemiddelde van een binomiale verdeling, gemiddeld 6 significante correlaties zijn gevonden.

$$\begin{aligned}
 M &= \frac{1}{2} \cdot n \cdot (n-1) \cdot P \\
 &= \frac{1}{2} \cdot 15 \cdot (15-1) \cdot P\{r_c \geq 0,30\} \\
 &= 105 \cdot 0,06 \\
 &= 6,3
 \end{aligned}$$

De overeenkomstindex (r_c) is 0,30. De kans, dat een nul-correlatie in een steekproef van 40 elementen groter of gelijk is aan 0,30, is 0,06 (zie bijlage I.1. geval 1a). Berekening van de standaard deviatie met behulp van de formule van de binomiale verdeling geeft:

$$\begin{aligned}
 S.D. &= \sqrt{\frac{1}{2} n \cdot (n-1) \cdot P \cdot Q} \\
 &= \sqrt{105 \cdot 0,06 \cdot 0,94} \\
 &= 2,4
 \end{aligned}$$

Bij analyse van groep 10 met een overeenkomstindex van 0,30 werden twee clusters gevonden: één bestaande uit elf operatoren en één uit twee operatoren. Dit betekent dat er $(11-1) + (2-1) = 11$ correlaties zijn gevonden groter of gelijk aan de overeenkomstindex. Het aantal operatoren van de clusters moet voor ieder cluster met één worden verminderd, omdat bij de keuze van de eerste operator geen vergelijking met de overeenkomstindex plaats vindt. De kans, dat bij aselechte data van 15 operatoren bij 40 elementen, 11 of meer significante correlaties voorkomen, is nu gelijk aan:

$$\begin{aligned}
 P\left\{t \geq \frac{11-M}{\sigma}\right\} &= P\left\{t \geq \frac{11-6,3}{2,4}\right\} = \\
 P\{t \geq 1,96\} &= 0,025
 \end{aligned}$$

De nulhypothese kan dus worden verworpen. De data van deze groep zijn niet aselekt en de clusters geen toevalsresultaat.

Bij bovengenoemde berekening is de benadering van de normaalverdeling voor de binomiaal gebruikt. Deze benadering is toegestaan, omdat M groter is dan 5 en P kleiner dan 0,5 (zie o.a. Hoel, 1962, p.109).

De overeenkomstindex van 0,30 voldoet aan de minimumwaarde van 15 operatoren bij een steekproefgrootte van 40 (zie tabel 4.6.1.). Wanneer de analyse met een overeenkomstindex van 0,25 was uitgevoerd en dus de index lager was dan de minimumwaarde nodig voor de toetsing, was de nulhypothese niet verworpen. In dat geval is $P\{r \geq 0,25\} = 0,096$ en dus $M = 10,08$ en $S.D. = 3,03$ en dus $P\{t \geq \frac{11-10,08}{3,3}\} = 0,889$. (De resultaten van de analyse van groep 10 zijn hetzelfde bij een overeenkomstindex van 0,25 of 0,30). Boven genoemd resultaat is het gevolg van een onvoldoende 'power' in dat geval van de toets. Het gemiddeld aantal 'significante' correlaties, dat dan bij toevallige resultaten kan voorkomen, is te hoog in vergelijking met het aantal dat maximaal kan worden gevonden. De formules, waarmee de overeenkomstindices van tabel 4.6.1 zijn berekend, stellen als eis dat dit gemiddelde aantal altijd kleiner is dan een $\frac{1}{2} n$.

Samenvattend kan worden gezegd dat controle van de clusters kan plaats vinden door

- a. het aantal operatoren van de cluster: dit moet voldoende groot zijn,
- b. bepaling van de gemeenschappelijke aspecten van de operatoren van het cluster in verschil met niet opgenomen operatoren: deze aspecten moeten relevant zijn, geen mogelijke niet-relevante condities betreffen,
- c. de K.R.(20) of (21) uitkomst moet significant zijn.

Verder dient te worden nagegaan

- d. of er niet te veel variatie in de ladingen van de clusteroperatoren is,
- e. bij *controle*-onderzoek of de correlaties met andere opgegeven

clusters volgens verwachting aan- of afwezig zijn, of bij *exploratief* onderzoek door inspectie van de correlaties van de niet-opgenomen operatoren of eventuele verstopte clusters aanwezig zijn, en

- f. vooral bij veel kleine clusters of het resultaat niet toevallig is (zie tabel 4.6.1.).

4.7. Rekenprogramma

Het programma van de bewerking heeft drie onderdelen,
te weten:

PREFAB; keuze van de uitgangsvariabele

ITERA ; de vorming van het cluster

OUTPUT; de berekeningen voor de output.

De input bestaat uit:

1. een nul en één matrix van operatoren ($1 \dots i \dots n$) en proefpersonen of stimuli ($1 \dots b \dots N$),
2. een overeenkomstindex: r_c per pivot,
3. pivots of sets van opgegeven operatoren ($1 \dots c' \dots C'$).

Bij *exploratie* omvat de pivot alle operatoren ($1 \text{ t/m } n$), bij *controle* meestal een gedeelte van de operatoren. In tabel 4.7.1. staat het blokdiagram van het programma. De daarbij gebruikte afkortingen hebben de volgende betekenis:

n ; totaal aantal operatoren,

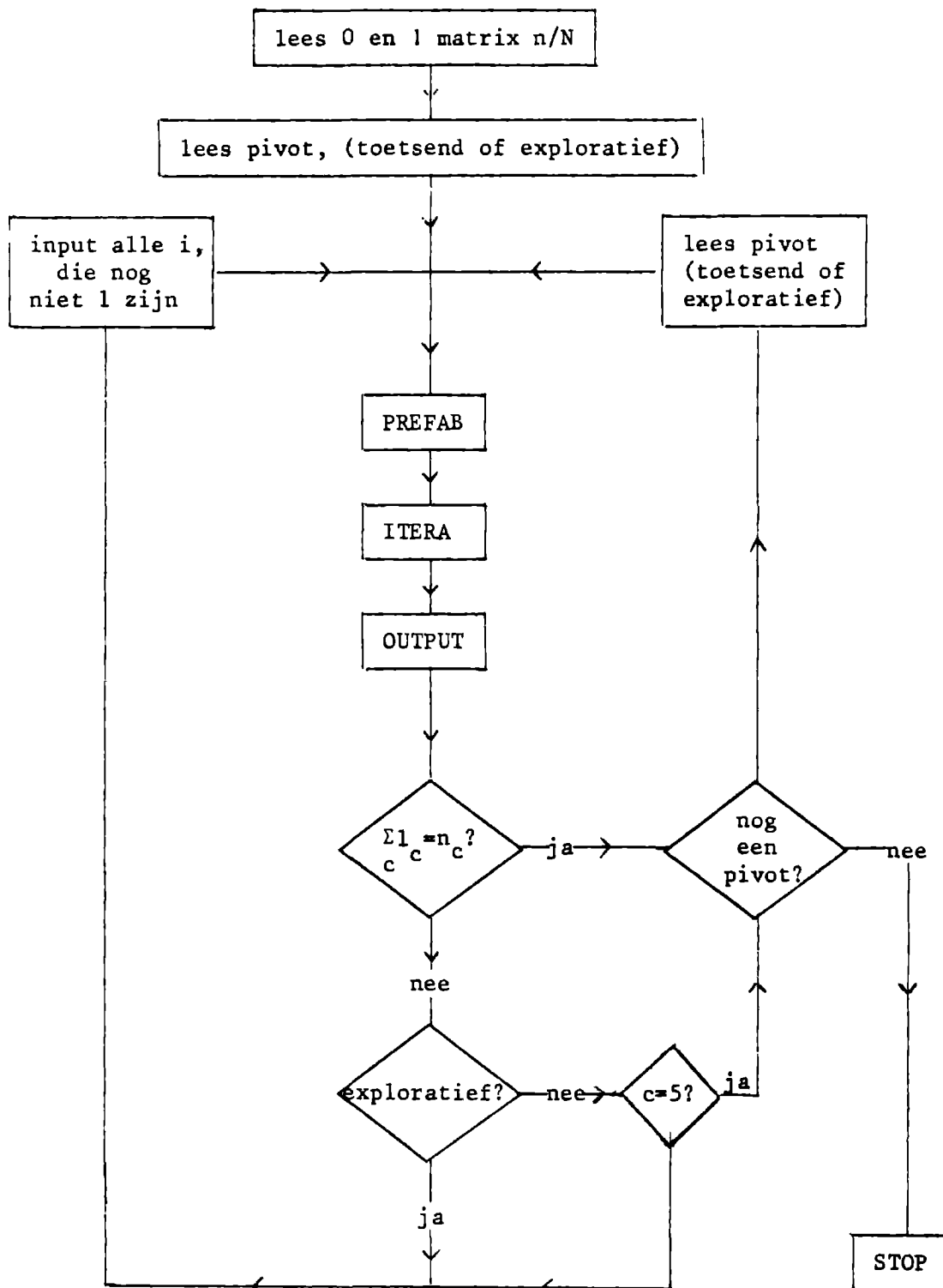
N ; totaal aantal proefpersonen of stimuli,

l ; het aantal operatoren opgenomen in de cluster,

c ; cluster ($1 \dots c \dots C$).

Tabel 4.7.1.

Blokdiagram hoofdprogramma



De onderdelen PREFAB, ITERA en OUTPUT geven elke keer één cluster.

Afhankelijk van het aantal operatoren in de clusters gaat het programma door of stopt. De stopregels zijn de volgende:

- a. wanneer alle operatoren i van de pivot zijn opgenomen in clusters $(\sum_c 1_c = n_c)$ en er geen pivot meer is, stopt het programma,
- b. wanneer het onderzoek niet exploratief is en er vijf clusters zijn gevonden ($c = 5$) en er geen pivot meer is, stopt het programma ook.

Bij de *controleerend-toetsende* bewerking worden er dus nooit meer dan vijf clusters per opgegeven pivot berekend. Deze beperking is om economische redenen ingevoerd. De input voor PREFAB is of een nieuwe pivot of de oude pivot met uitschakeling van de operatoren van de gevonden clusters. Deze input kan afwijken van die van ITERA. Bij de *controleerend-toetsende* bewerking vormen ook de operatoren van de gevonden clusters de input.

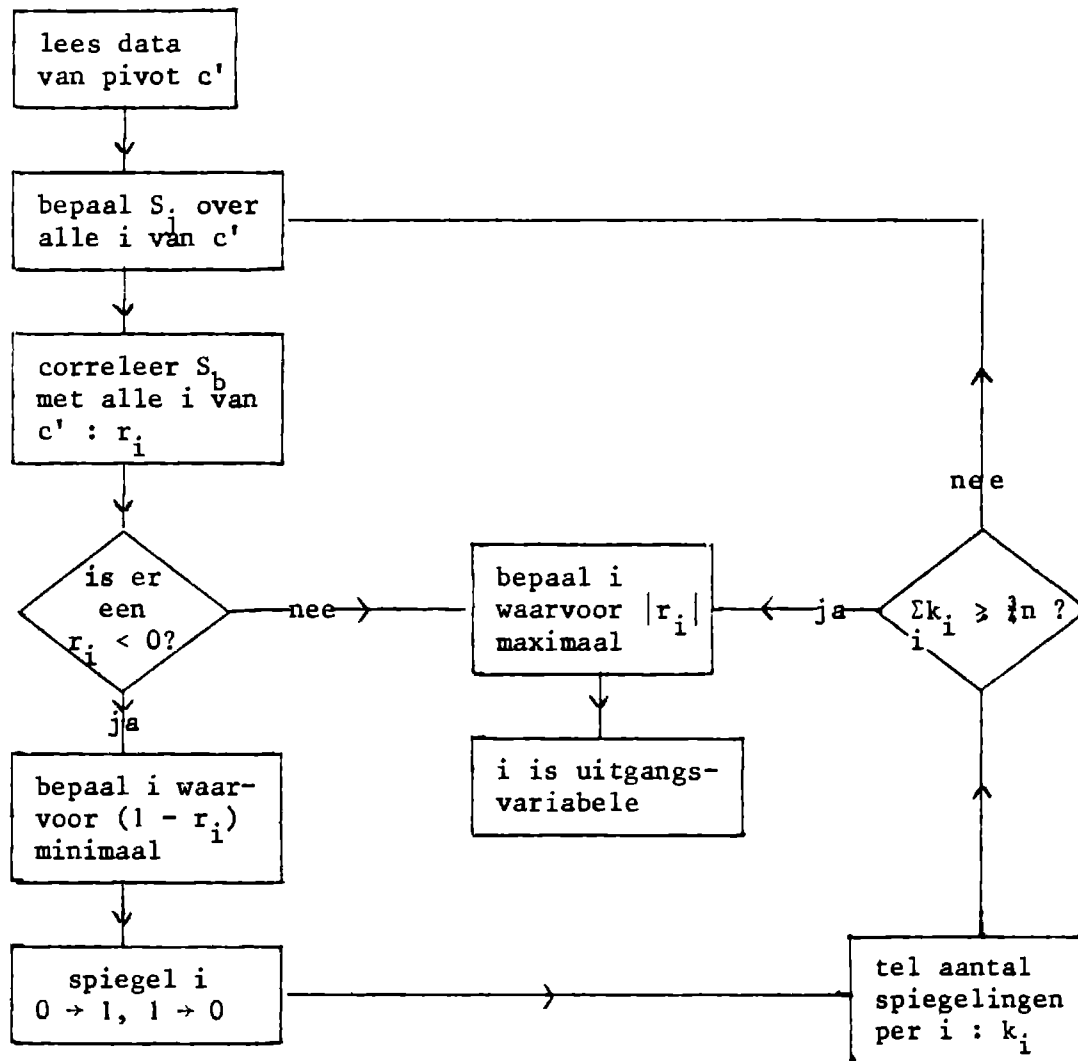
In tabel 4.7.2. staat het blokdiagram van PREFAB. De gebruikte afkortingen hebben de volgende betekenis:

- S_b ; de somscore van proefpersoon of stimulus b ,
 r_i ; de correlatie van S met operator i ,
 k_i ; aantal spiegelingen van operator i .

Tabel 4.7.2.

Blokdiagram PREFAB

(bepaling uitgangsvariabele)



Dit programma gaat door met spiegelen van operatoren zolang als er één correlatie negatief is ($r_i < 0$) of wanneer het totaal aantal maal dat er gespiegeld is ($\sum_i k_i$) gelijk of groter is dan driekwart van het totaal aantal operatoren.

De formule voor berekening van de correlatie r is de bekende productmoment formule

$$r = \frac{N \sum t \cdot S - (\sum t)(\sum S)}{\sqrt{[N \sum t^2 - (\sum t)^2][N \sum S^2 - (\sum S)^2]}}$$

Hierbij is t de uitslag van de operator ($t = 0$ of $t = 1$),

S de somscore, en

N het totaal aantal proefpersonen of stimuli b .

Alle summaties vinden plaats over $1 \dots b \dots N$.

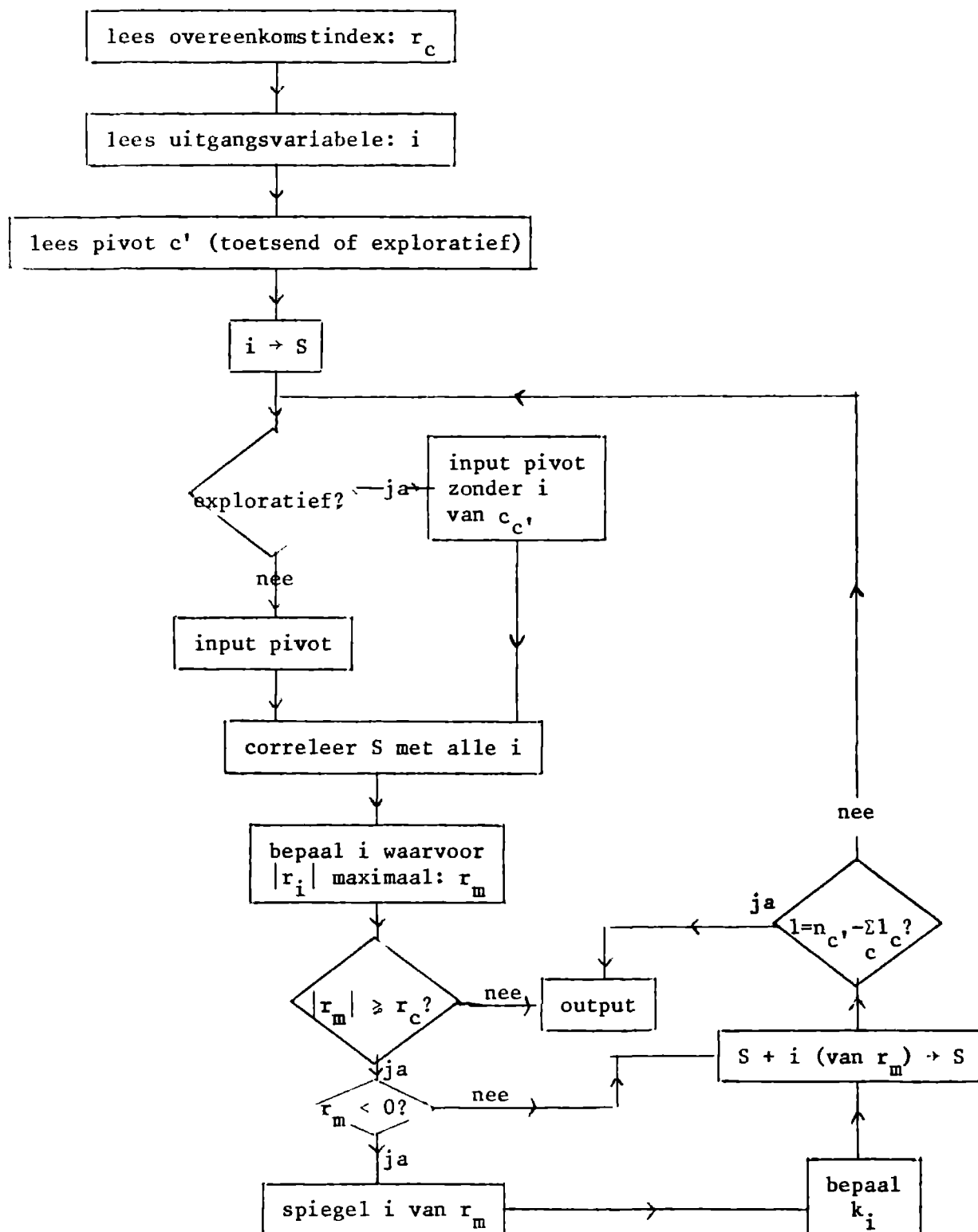
In tabel 4.7.3. staat het blokdiagram van ITERA. De afkortingen hebben de volgende betekenis:

r_m ; de maximale correlatie van alle r_i bij deze S ,

c_c ; de clusters van de pivot c' .

Tabel 4.7.3.

Blokdiagram van ITERA
(bepaling cluster)



De cluster wordt dus uitgebreid met een operator, zolang de absolute waarde van de correlatie maar gelijk of groter is dan de overeenkomst-index ($|r_m| \geq r_c$). Voor de opname van de nieuwe operator moet nog eerst worden nagegaan of hij niet gespiegeld moet worden ($r_m < 0?$). De input is bij *controlerend-toetsende* bewerking de gehele pivot, bij de *controlerend-exploratieve* of *exploratieve* bewerking de pivot met uitsluiting van de al in clusters van de pivot opgenomen variabelen.

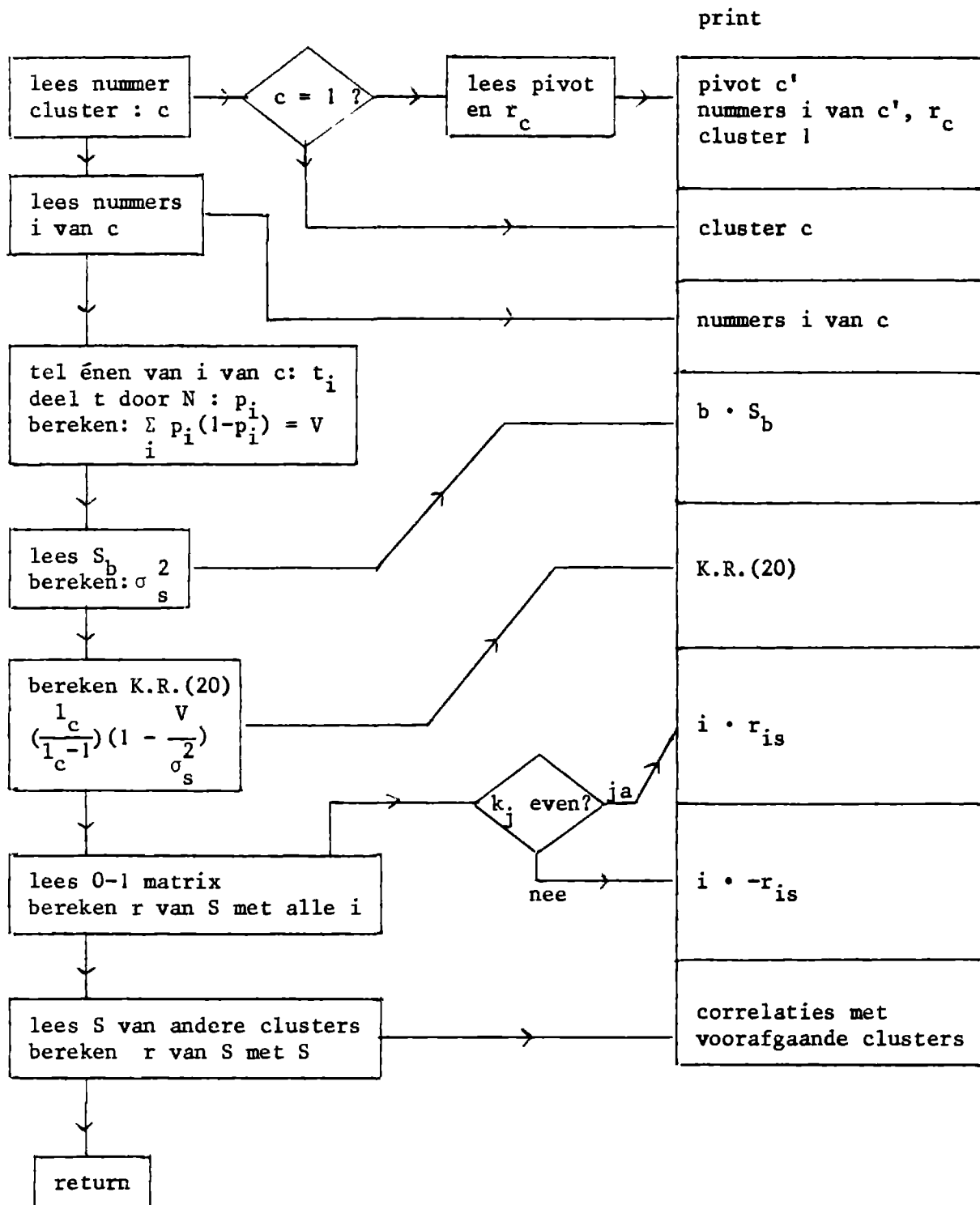
In tabel 4.7.4. staat het blokdiagram van OUTPUT. Als print wordt gegeven:

1. bij de eerste cluster de pivot en zijn overeenkomstindex,
2. het nummer van de cluster,
3. de operatoren van de cluster,
4. de clusterscores (S_j) per proefpersoon of stimulus,
5. de K.R.(20),
6. de correlaties van S met *alle* operatoren ook van andere pivots, en
7. de correlaties van S met de S van andere clusters.

In tabel 4.7.5. staat een voorbeeld van een mogelijke print. De tabel spreekt verder voor zichzelf.

Tabel 4.7.4.

Blokdiagram OUTPUT



Tabel 4.7.5.

Voorbeeld van een print

| | |
|----------------------------|------------------------------|
| <u>pivot II</u> | <u>pivot c'</u> |
| 1 5 6 7 9 10 21 | |
| 31 32 34 40 42 51 54 | nummers i van c' |
| 62 64 65 66 71 75 8 | |
| 95 96 97 98 | |
| 0,30 | overeenkomstindex |
| <u>cluster 1</u> | <u>cluster c</u> |
| 75 95 96 97 98 | |
| 21 71 54 59 69 | nummers i van c |
| 1.7 2.11 3.16 4.15 | |
| 31.15 32.6 33.8 34.2 | clusterscores |
| 61.4 62.12 63.7 64.10 | b S_b |
| K.R.(20) = 0,868 | K.R.(20) |
| 1. 0,044 2. 0,219 | ladingen (onderstreept) en |
| 11. -0,034 12. 0,003 | correlaties met alle andere |
| 21. <u>0,465</u> 22. 0,078 | operatoren |
| 31. -0,069 32. 0,0 | $i \cdot r_{is}$ |
| 41. <u>0,486</u> 42. 0, | |
| 51. 0,278 52. | |
| 61. <u>0,471</u> 6 | |
| 71. <u>-0,510</u> | |
| 81. 0,395 | |
| 91. 0,002 | |
| I 1 0,03 | correlaties met voorafgaande |
| I 2 -0,16 | clusters |

Bovenstaande rekenprogramma kan met een rekenmachine worden uitgevoerd op data met een n/N , die niet groter is dan 100 bij 200. Het is mogelijk een aantal vereenvoudigingen in te voeren onder andere door niet volledig in de verschillende stadia de correlaties uit te rekenen, berekening van de covariantie is meestal voldoende voor het nemen van de beslissingen.

Sneller en nagenoeg foutloos vindt de berekening plaats wanneer een computerprogramma wordt gebruikt. De bewerking kan dan op veel meer data worden uitgevoerd, onder andere doordat overbelasting van het geheugen kan worden voorkomen door de 0/1 data met een UTMOST-programma op te bergen in de binaire posities van het geheugen. Hierdoor wordt het mogelijk de bewerking uit te voeren, afhankelijk van de computer, op data met een n/N van 1000 bij 1000. Voor de berekening van de correlaties tussen de clusters moeten dan wel de somscores van de gevonden clusters worden bewaard. Dit kan problemen opleveren, vooral omdat vaak niet bekend is hoeveel clusters gevonden zullen worden. Een oplossing is mogelijk door de somscores op tape te zetten en pas aan het eind van de analyse de correlaties te berekenen.

De snelheid van het computerprogramma is voornamelijk afhankelijk van het aantal operatoren per pivot, het aantal pivots en het aantal clusters, dat wordt gevonden. Hoe meer clusters, hoe langer de bewerking duurt en wel voornamelijk door het vele printwerk. Het verdient daarom aanbeveling 'stoppen' in te voeren, als bijv. bij exploratie niet meer dan $1/5 i_c$ clusters of bij toetsing niet meer dan 5. Gemiddelde tijden zijn voor bijv. 14 operatoren per pivot 90", 100 operatoren per pivot 20'. Veel hangt ook af van de

wijze waarop het programma is geschreven. Het is bijv. mogelijk tijdens de print van een cluster door te gaan met de berekeningen voor het volgende cluster. Een verdere tijdsbesparing kan worden gevonden door alleen de correlaties uit te rekenen bij de output; het is niet nodig voor spiegeling of bepaling maximum de correlatie volledig te berekenen. Het is verder niet zo dat de gegeven blok-diagrammen een optimaal computerprogramma geven. Er kunnen op verschillende wijzen vereenvoudigingen worden ingevoerd. C.H.J. Schreibers en H.E.M. Borgers hebben ieder een dergelijk programma geschreven voor de iteratieve clusteranalyse. De eerste in FORTRAN voor de UNIVAC III van de N.V. Nederlandse Staatsmijnen (Heerlen). De tweede in PLI voor de IBM 360-50 van de Katholieke Universiteit (Nijmegen). Bij het maken van een computerprogramma verdient het aanbeveling contact op te nemen met één van deze programmeurs.

5. Toepassingen van de analyse

5.1. Inleiding

Voor de toepassing van de analyse dient de onderzoeker een aantal beslissingen te nemen ten aanzien van de uitvoering. De analyse kan, zoals al is duidelijk gemaakt, controlerend-toetsend, controlerend-exploratief of explorerend worden uitgevoerd. Toepassing van de analyse met verschillende overeenkomstindices geeft de mogelijkheid de orde van de klassificatie te onderzoeken. Daarnaast kan de analyse als R-analyse worden uitgevoerd of als Q-analyse. In het eerste geval gaat het om klassificaties van operatoren of kenmerken en in het tweede geval om klassificaties van elementen.

R- en Q-analyse van de data, hoewel principieel overeenkomstig, geven verschillende informatie over de data. De beslissing hierover hangt af van de *soort van toepassing*. Controlerend of explorerend onderzoek is afhankelijk van de kennis, die aanwezig is over het betreffende gebied. Des te meer de onderzoeker in staat is een conditioneel model voor de data op te stellen, des te meer kan hij controlerend te werk gaan. Het gaat hier om de *vorm van toepassing*. Hetzelfde geldt voor onderzoek naar de hiërarchie of orde van de klassificatie en predictie-onderzoek.

Allereerst zullen de verschillende soorten van toepassingen worden behandeld, daarna zal op de vormen worden ingegaan.

5.2. Soorten van toepassing

Om tot een adequate toepassing van de analyse te komen, zal de onderzoeker zich moeten realiseren wat het doel van het onderzoek is in het kader van de in de hoofdstukken 2. en 3. ontwikkelde theorie.

Is het doel van het onderzoek informatie te verkrijgen over 'stimuli' of over 'individueen'? In de praktijk blijkt het antwoord vaak moeilijkheden op te leveren, omdat informatie over 'stimuli' vaak verkregen wordt met behulp van beoordelingen of uitspraken van personen. Meestal zijn deze personen echter onderdelen van de operatoren en geen 'proefpersonen', zoals bij tests het geval is (zie ook paragraaf 2.3.). De items van tests zijn ook op te vatten als 'stimuli', als elementen, die dragers zijn van bepaalde eigenschappen. Toch gaat het bij testconstructie om informatie over 'individueen', om differentiërende eigenschappen van mensen of dieren. Een duidelijke uitspraak op bovengenoemde vraag is mogelijk door na te gaan welke veranderingen met behulp van de informatie van het onderzoek aangebracht kunnen worden. Wanneer deze veranderingen de 'stimuli' betreffen, zoals bij leerprogramma's, functie-inhouden, woonplaatsen, televisieprogramma's, produkten, goederen, werksituaties, organisaties het geval is, dan dienen de elementen van het onderzoek 'stimuli' te zijn. Gaat het echter om veranderingen van individuen, zoals bij om-, her-, bijscholing, counseling, school- of beroepskeuze en selectie het geval is, dan zijn de elementen 'individueen'. Het verschil correspondeert grotendeels met de objecten van respectievelijk de algemene psychologie en de differentiële psychologie. Een formeel criterium wordt gegeven door de vraag of de gebruikte 'stimuli' vallen onder de definitie van de operator, zoals gesteld in paragraaf 2.3. In dat geval gaat het onderzoek over 'personen'. Wanneer echter de 'personen' vallen onder de genoemde definitie, dan zijn de 'stimuli' de elementen van het onderzoek.

Een tweede uitspraak is nodig om te bepalen of R- dan wel Q-analyse moet worden uitgevoerd. Sinds de studie van Broverman (1961) is de reden van de controversie tussen Stephenson (1953) en Cattell (1952) over R- en Q-analyse duidelijker geworden. Beide analyses geven principieel dezelfde informatie. De clusters van R-analyse, waar operatoren op hun overeenkomst worden geklassificeerd, zijn directe transposities van de clusters van Q-analyse, waar de elementen worden geklassificeerd, onder voorwaarde, dat beide analyses uitgaan van een zelfde dubbelgestandaardiseerde (voor zowel de rijen als de kolommen) datamatrix. Dit laatste is nu meestal niet het geval en daarom geven beide analyses vaak een andere presentatie van de data. Voor een nul en één matrix (de data van de iteratieve clusteranalyse) betekent dit dat de p-waarden (de kans op code één) van de rijen gelijk dienen te zijn evenals de p-waarden van de kolommen. Wanneer dit het geval is, is een uitspraak over R- en Q-analyse niet relevant. Beide analyses geven dezelfde weliswaar op verschillende wijze gepresenteerde informatie, namelijk zodanig dat de ladingen transposities zijn van de scores en omgekeerd.

R-analyse geeft een klassificatie van de operatoren of kenmerken en maten voor de elementen. Q-analyse geeft een klassificatie van de elementen en maten voor de kenmerken. Een beslissing over R- of Q-analyse is mogelijk door na te gaan in hoeverre de voorwaarde van gelijke p-waarden acceptabeler is voor de kolommen (operatoren) of voor de rijen (elementen). In het eerste geval is R-analyse de aangewezen methode, in het tweede Q. De vraag wordt dan: kunnen operatoren of juist elementen met extreme p-waarden uit het onderzoek worden gemist? Bij prestatietests, vragenlijsten of beoordelingsschalen is vaak het eerste het geval items met zeer grote of zeer kleine moeilijkheidsgraad of populariteitsgraad zijn vaak niet relevant en dus is R-analyse de aangewezen methode. Individuen of proefpersonen met zeer hoge of lage p-waarden kunnen echter niet worden weggelaten, omdat het om

interindividueel onderzoek of vergelijkingen tussen mensen gaat.

Bij projectietests is echter een weinig voorkomend antwoord soms zeer relevant en moet de testafname bij proefpersonen, die geen of nagenoeg geen antwoorden geven, als mislukt worden beschouwd. Het gaat namelijk dan niet om niveau-bepalingen van de proefpersoon voor wat betreft bijvoorbeeld zijn creativiteit, maar om diagnose van zijn belevingswereld. Het gaat om typering van belevingen, om *type-onderzoek*. Q-analyse is dus aangewezen. Bij *betekenisonderzoek* van stimuli kunnen ook geen kenmerken worden gemist, terwijl bij onderzoek van *psychische functies* juist stimuli met sterk afwijkende p-waarden gehandhaafd dienen te worden. Het eerste onderzoek vraagt dus om Q-analyse, het tweede om R-analyse.

Het in tabel 5.2.1. gegeven schema vat bovenstaande samen.

Tabel 5.2.1.

Uitvoeringsschema van de analyse afhankelijk
van het doel van het onderzoek

| elementen | R-analyse | Q-analyse |
|---------------|-------------------------------|--------------------|
| stimuli | psychische functies | betekenisonderzoek |
| proefpersonen | interindividueel onderzoek | type-onderzoek |

5.2.1. Betekenisonderzoek

Eén van de meest bekende voorbeelden van betekenisonderzoek is het gebruik van de *Semantische Differentiaal* van Osgood (zie Nederlands Instituut van Psychologen, 19.2., p. 220). Het doel van deze meet-schaal is de bepaling van de affectieve betekenis van begrippen. Met de clusteranalyse kan nu worden nagegaan hoe de klassificatie is van een omschreven groep begrippen op grond van hun gemiddelde beoordelingen per factor of per polariteitsschaal van een groep beoordelaars.

Het is daarbij niet nodig zich te beperken tot 'begrippen'. Ook 'beroepen', 'loopbanen', 'wegen', 'landen', 'woorden', kunnen op deze wijze worden geklassificeerd. Belangrijk hierbij is op grond van welke gezichtspunten of kenmerken de klassificatie moet plaats vinden. Het kan gaan om de beroepseisen, de mate van succes van loopbanen, de veiligheid van wegen, het nationaal bewustzijn van landen, de leesbaarheid van woorden. Voor deze kenmerken worden indices, beoordelingen of antwoorden verzameld en voor iedere stimulus moet zoveel mogelijk een gelijk aantal van deze kenmerken opgaan.

Dergelijke onderzoeken hebben meestal tot doel te komen tot een indeling of klassering van een groot aantal verschillende stimuli op grond van hun kenmerken. Hierdoor wordt een interpreteerbaar overzicht verkregen van de stimuli. Om die reden kunnen stimuli, waar nagenoeg alle kenmerken voor opgaan, of stimuli, waar juist geen enkel kenmerk op van toepassing is, buiten het onderzoek worden gelaten. Beide groepen vormen al een duidelijk onderscheidbare en ook interpreteerbare klasse. Voor de stimuli, waar gemiddeld 50% van de kenmerken opgaan, is echter een indeling op frequentie van het aantal kenmerken weinig doelmatig. Er moet dan gezocht worden naar groepen, die bepaalde kenmerken wél en andere kenmerken niet gemeenschappelijk hebben, dat wil zeggen naar clusters van stimuli.

De gegeven uitzettingen kunnen nader worden geïllustreerd met een analyse van beroepen van leerlingen van huishoudscholen. Dit onderzoek, dat is uitgevoerd door Scheerder en Kerkhoff (1969) betreft de analyse van een lijst van 47 beroepen opgesteld door beroepsdeskundigen en directrices van huishoudscholen als toegankelijk voor L.H.N.O.meisjes.

Deze beroepen zijn op achttien punten beoordeeld aan de hand van de eisen van de beroepsbeelden van de Documentatie Opleidings- en Studiemogelijkheden van het Rijksarbeidsbureau (Min.v.Soc.Zaken en Volksgez., 1963). De veertien beroepseisen van deze publicatie zijn de volgende.

1. *Exact*: de aanleg, welke gericht is op het *scherp en nauwkeurig onderkennen* van gegeven feitelijkheden: het toetsen van deze feitelijkheden al naar gelang van hun aard aan natuurwettelijke of door de mens gestelde namen; het oplossen van ten aanzien van de gegeven feitelijkheden, gestelde problemen, of op het vaststellen van de doelmatige toepassingsmogelijkheden en uitvoeringswijzen. De feitelijkheden kunnen zijn eigenschappen van stoffen, hoeveelheden, cijfermateriaal, etc.

B.v. actuaris - apothekersassistent - boer - boekhouder - advocaat - technisch tekenaar.

2. *Technisch*: de aanleg, welke gericht is op het doelmatig *vormgeven* aan vaste stoffen en op het werken met deze doelmatig gevormde stoffen. Het gaat hier om het ontwerpen van de doelmatige vorm voor gebruiksvoorwerpen, machines, etc. om het vaststellen van de wijze van *vervaardiging* en uitvoering en tevens om de uitvoering zelf.

B.v. metselaar - kleermaker - schoenmaker.

3. *Organisatorisch*: de aanleg, welke gericht is op het *ordenen van menselijke arbeid* en materiële middelen, als regel productiemiddelen,

tot een doelmatig functionerend geheel.

B.v. bedrijfsleider - boer - reisgids.

4. *Verbaal*: de aanleg, welke er op gericht is om *gedachten* mondeling of schriftelijk op een bevattelijke wijze te vertolken en om anderzijds te begrijpen hetgeen anderen hebben gesproken of geschreven.

B.v. machinezetter - stenotypiste - schrijver - leraar.

5. *Sociaal*: de aanleg, welke er op gericht is om taktvol en in een *goede* verstandhouding met mensen om te gaan en tevens aan een *goede verstandhouding tussen derden* onderling te bevorderen en tot stand te brengen. Het gaat hier om *begrip* en gevoel voor de onderlinge verhoudingen en het verkeer tussen mensen. Dit om zichzelf en anderen welke men heeft te leiden en te richten, aan de samenleving aan te passen met respectering van de persoonlijkheid van anderen en tevens met handhaving van eigen persoonlijkheid en gezag.

B.v. kinderverzorgster - onderwijzer - psycholoog - politie-agent.

6. *Kunstzinnig*: de aanleg, welke er op gericht is om een beleving op een zodanige wijze te uiten, dat hierdoor bij anderen een analoge gevoelsstroom ontstaan.

B.v. bloemenbinder - beeldhouwer - schrijver.

7. *Materiaalgevoel*: de aanleg, welke er op gericht is om op grond van *waarnemingen* en voorafgaande ervaring, *aan te voelen* welke toepassings- en bewerkingsmogelijkheden materialen bieden en hoe gegeven materialen *moeten worden behandeld* en bewerkt.

B.v. metaaldraaier - broodbakker - kok.

8. *Vormgevoel*: de aanleg, welke er op gericht is, om de vorm en omtrek van voorwerpen scherp in zich op te nemen en in zijn geest

voor te stellen en om de juiste doelmatige vorm voor produkten te *bepalen*.

B.v. hoedenmaker - schoenmaker - beeldhouwer.

9. *Aandacht*: de aanleg, welke er op gericht is, om de gedachten *sterk* op het werk te richten en gericht te houden. Het staat primair bij beroepen, waar het gaat om het sorteren, nakijken en tellen van produkten.

B.v. drukwerkkontroleur - chauffeur - machine-inpakker.

10. *Nauwkeurigheid*: de aanleg, welke er op gericht is om werkzaamheden te verrichten welke aan *enige vorm van precisie* zijn gebonden. Deze precisie kan betrekking hebben op afmetingen, de juiste plaats e.c. Het wordt hier alleen gebruikt als nadere bepaling van handvaardigheid.

B.v. schoenstikster - knoopsgatenmaker - horlogemaker.

11. *Toewijding*: de aanleg, welke er op gericht is om zich met *zelf-overgave* op het werk toe te leggen.

B.v. verpleegster - onderwijzer - dienstbode - dierenverpleger.

12. *Handvaardigheid*: de aanleg, welke er op gericht is om de bewegingen en verrichtingen van de *handen* zodanig te besturen als door het werk wordt vereist. Het gaat hier om *meer* dan gewone handvaardigheid, welke niet door iedereen kan worden opgebracht.

B.v. graveur - kleermaker - boetseerder - chirurg..

13. *Persoonlijk voorkomen*: de uiterlijke verschijning en de aanleg om door korrekt en gereserveerd optreden tegenover publiek en klanten een *aangename* indruk te maken. Het verschil met sociaal is, dat het hier *niet* gaat om een nauwe omgang en verkeer met mensen, maar slechts om een *oppervlakkig* en

voorbijgaand kontakt.

B.v. kantoorbode - liftbediende - kondukteur - kelner.

14. *Fysieke inspanning*: beroepen, welke tamelijk veel lichamelijke inspanning vereisen.

Aan de 14 eisen waren nog 4 toegevoegd, nl.:

15. *Handvaardigheid* belangrijker dan *toewijding*.

16. *Handvaardigheid* belangrijker dan *naauwkeurigheid*.

17. *Handvaardigheid* belangrijker dan *aandacht*.

18. *Vormgeving* belangrijker dan *handvaardigheid*.

De beoordelingen werden op een niet-systematische wijze gegeven door een panel van beroepsdeskundigen onder leiding van L. van Geffen (G.I.T.P. Nijmegen en Psychol. Lab. Afd. Bedrijfspsychol.). In bijlage II.3. staat de nul en één matrix van dit onderzoek. Eén betekent de beroepseis is relevant, nul niet-relevant. Voor ieder beroep zijn 4 of 3 eisen aangegeven. Dit gaf geen problemen gezien het nagenoeg gelijke vooropleidingsniveau van de beroepen (beroepen uit de eerste vier klassen van de beroepsklassificatie).

Vanwege het doel van het onderzoek (klassificatie van beroepen) is een analyse over de oordelen (Q-analyse) het meest aangewezen. De mogelijkheden voor informatie zijn dan ook goed: 47 over 18. De statistische betrouwbaarheid geeft echter, gezien de steekproefgrootte van 18, problemen. Ernstig moet rekening worden gehouden met kunstmatige correlaties bij een zo'n kleine steekproefgrootte. Tabel 4.6.1. leert dat alleen als een gemiddelde intercorrelatie van 0,60 tussen clustervariabelen wordt verwacht deze analyse zinvol uitvoerbaar is. Gezien de grote aandacht, die aan de oordelen is gegeven, is dit echter denkkelijk geen bezwaar. De analyse heeft dus plaats

gevonden over de 18 oordelen met een overeenkomstindex van 0,60.

De analyse geeft 22 clusters, waarvan er 14 specifiek zijn. De K.R. (21) waarden van de 8 niet-specifieke clusters zijn alle significant.

Totaal zijn 25 'significante' correlaties gevonden. Dit resultaat is duidelijk niet toevallig. Bij een toevalsresultaat zouden gemiddeld 9 'significante' correlaties zijn gevonden met een gemiddelde standaardafwijking van 3. De berekeningen hiervoor zijn als volgt (zie paragraaf 4.6.):

$$t' = \frac{0,60 \sqrt{18-2}}{\sqrt{1-0,36}} = 3,0$$

$$P \{ t' \geq |3,0| \} = 0,008$$

$$M = \frac{1}{2} \cdot 47 \cdot (46-1) \cdot 0,008 = 9$$

$$S.D. = \sqrt{\frac{1}{2} \cdot 47 \cdot (46-1) \cdot 0,008 \cdot 0,992} = 3$$

We geven hier de gevonden clusters van de analyse met een overeenkomstindex van 0,50. Gezien het resultaat van bovenstaande toetsing is daar geen bezwaar tegen. Bij 16 vrijheidsgraden is een correlatie van 0,50 significant voor 3½%. De clusters worden in dit geval beter gepresenteerd. Ieder cluster wordt beschreven met zijn beroepen, achter ieder beroep staat de lading van dit beroep in het betreffende cluster. Tevens worden de relevante eisen genoemd. Achter ieder genoemde eis staat de klassificatiemaat van deze eis. De klassificatiematen geven aan in welke mate de eisen relevant zijn voor de beroepen van het cluster. De maximaal mogelijke klassificatiemaat is gelijk aan één. Een klassificatiemaat is namelijk het aantal keren dat de eis relevant is voor de beroepen van het cluster (de clusterscore) gedeeld door het totale aantal beroepen van het cluster (zie paragraaf 3.3.). Weggelaten zijn de eisen met lagere maten dan de genoemde, evenals de beroepen, die niet in het cluster zijn opgenomen.

cluster 1(Q) K.R. (21) = 0,90

- 13. etaleuse, 0,93
- 18. hoediennaaiester, 0,87
- 5. bloemschikster, 0,76
- 25. kapster, 0,76
- 35. pottenbakster, 0,76
- 29. modiste, 0,76
- 20. kostuumnaaister, 0,75
- 6. kunstzinnig, 0,86
- 8. vormgevoel, 0,86
- 18. vormgeving belangrijker dan handvaardigheid, 0,86

cluster 2(Q) K.R. (21) = 0,92

- 2. assistent-groepsleidster, 0,92
- 12. dokters-assistente, 0,91
- 37. serveerster, 0,91
- 1. assistent-jeugdleidster, 0,80
- 11. dierenarts-assistente, 0,80
- 8. kassière (ook in cluster 3), 0,74
- 9. konduktrice (ook in cluster 7), 0,63
- 4. verbaal, 1,00
- 13. persoonlijk voorkomen, 1,00

cluster 3(Q) K.R. (21) = 0,87

- 7. controleuse in de industrie (ook in cluster 7), 0,87
- 26. kantoor-assistente, 0,80
- 31. modinette, 0,80
- 38. schoenstikster, 0,80
- 42. telefoniste, 0,73
- 32. nopster-stopster, 0,66

8. kassière (ook in cluster 2), 0,53

9. nauwkeurigheid, 1,00

10. aandacht, 1,00

cluster 1(1) K.R. (21) = 0,93

22. kinderverzorgster, 0,89

23. kleuterhulp, 0,89

44. verpleegster (in psych. inrichting), 0,89

46. ziekenverzorgster, 0,89

4. bejaardenverzorgster, 0,81

19. inrichtings-assistente, 0,79

16. huishoudhulp, 0,69

47. zweminstructrice, 0,69

11. toewijding, 1,00

14. fysieke inspanning, 1,00

cluster 5(1) K.R. (21) = 0,86

30. monteuse in bedrijf, 0,89

36. ponster en/of kodeuse, 0,89

41. tricotnaaister, 0,89

24. kamermeisje, 0,40

2. technisch, 0,75

12. handvaardigheid, 0,75

15. handvaardigheid belangrijker dan toewijding, 0,75

cluster 6(1) K.R. (21) = 0,83

14. gezinshulp-verzorgster, 0,93

21. kraamverzorgster (ook in cluster 8), 0,89

40. strijkster, 0,78

3. ateliernaaiester, 0,46
33. pakster, 0,46
14. fysieke inspanning, 0,60
15. handvaardigheid belangrijker dan toewijding, 0,60
17. handvaardigheid belangrijker dan aandacht, 0,60
- cluster 7(Q)* K.R. (21) = 0,71
43. tandartsassistente, 0,87
45. verkoopster, 0,87
9. konduktrice (ook in cluster 2), 0,52
7. kontroleuse in de industrie (ook in cluster 3),
0,52
1. exact, 0,50
2. technisch, 0,50
13. persoonlijk voorkomen, 0,50
- cluster 8(Q)* K.R. (21) = 0,81
15. huisvrouw, 0,92
17. huisjuffrouw, 0,92
21. kraamverzorgster (ook in cluster 6), 0,45
3. organisatorisch, 0,67
11. toewijding, 0,67
17. handvaardigheid belangrijker dan aandacht, 0,67

Specifieke clusters werden gevormd door

39. schoonheidsspecialiste
27. kookster - kokkin
6. confectionaaiester
28. linnerjuffrouw
10. chauffeuse

De clusters van beroepen spreken voor zich. Daarom zullen hier verder geen interpretaties worden gegeven.

Bij betekenisonderzoek wordt er vaak van uitgegaan, dat de stimuli adequaat beschreven zijn met het gemiddelde of op een andere wijze samengesteld oordeel van een groep deskundigen of personen. Het is wel nodig voor deze aanname te waarschuwen. Uit toepassingen van de clusteranalyse bij enquête-onderzoek, waar een 'situatie' of complex van stimuli worden beoordeeld, is gebleken, dat dit lang niet altijd het geval is (zie paragraaf 5.2.4.).

5.2.2. Psychische functies

Om bepaalde cognitieve, affectieve of conatieve functies bij de mens of dier te bestuderen is het mogelijk individuen verschillende stimuli aan te bieden en dan na te gaan hoe deze stimuli worden verwerkt. In dat geval zijn de stimuli de elementen en is de geobserveerde en gecodeerde reactie van de individuen de operator. De analyse dient dan echter als R-analyse over deze elementen te worden uitgevoerd. Reacties namelijk, die voor alle stimuli opgaan, kunnen dan buiten de analyse worden gehouden. Deze reacties zijn niet differentierend binnen de groep aangeboden stimuli. Stimuli, die echter een groot aantal reacties ontlokken of juist geen, dienen te worden gehandhaafd.

Ook bij deze onderzoekingen is het belangrijk het gezichtspunt vast te stellen van waaruit de reacties worden geobserveerd en gecodeerd. Bij bestudering van cognitieve processen zullen daarom vooral opdrachten met de stimuli van belang zijn. Bij affectieve functies echter juist de belevingen en waarderingen van de stimuli, terwijl studie van conatieve functies vraagt om een benadering uit de behoefte sfeer door de beoordelaars.

De analyse gaat dan na welke reacties als overeenkomstig kunnen worden beschouwd en welke stimuli vooral deze reacties oproepen. De analyse zal dan echter wel, evenals bij betekenisonderzoek, moeten werken met gemiddelde reacties van beoordelaars. Analyse van een drie-dimensionele datamatrix is (nog) niet mogelijk. In de vorige paragraaf is al op de problemen van gebruik van gemiddelden en hun mogelijke oplossing gewezen. Wanneer maar één reactie van de beoordelaars wordt gevraagd kan het onderzoek ook beschouwd worden als type-onderzoek. Een dergelijk onderzoek wordt dan uitgevoerd als Q-analyse met de beoordelaars als elementen.

De drie factoren van Osgood - evaluatie, potentie en dynamiek - zijn een voorbeeld van een resultaat bij toepassing van klassificatiemethoden in het affectieve gebied. Het halo-effect van personeelsbeoordelingen is ook een factor of cluster, die bij R-analyse van stimuli (de beoordeelde) wordt gevonden. Bovenstaande is nog nader te illustreren door de datamatrix van de beroepen uit de vorige paragraaf (zie ook bijlage II.3.) opnieuw te analyseren maar nu met een R-analyse.

De p-waarden van de beroepseisen lopen ook niet veel uiteen. De gemiddelde p-waarde is 0,22 met een range van 0,17 t/m 0,28. Verwacht mag dus worden, dat de resultaten van de R- en Q-analyse van deze datamatrix elkaar niet veel zullen ontlopen. Bij uitvoering als R-analyse wordt de datamatrix 18 bij 47. Hierdoor kan met een lagere overeenkomstindex dan 0,60 worden gewerkt. Een overeenkomstindex van 0,60 gaf maar één niet-specifiek cluster namelijk van 'vormgeving' (8), 'kunstzinnig' (6) en 'vormgeving belangrijker dan handvaardigheid' (18), en wel overeenkomstig het eerste cluster van de Q-analyse. De analyse heeft plaats gevonden met een overeenkomstindex van 0,37 (signifcant voor 1%, $df=45$). Hiermee werden 9 clusters gevonden, waarvan 6 specifiek en wel

| | |
|----------------------|-------------------------------------------------|
| <u>cluster 1 (R)</u> | K.R. (21) = 0,62 |
| 14. | fysieke inspanning, 0,77 |
| 11. | toewijding, 0,71 |
| 5. | sociaal, 0,66 |
| 16. | handvaardigheid belangrijker dan aandacht, 0,58 |
| 10. | nauwkeurigheid, -0,66 |
| 9. | aandacht, -0,62 |
| 1. | exact, -0,58 |

- 46. ziekenverzorgster, 1,00
- 19. inrichtingsassistente, 1,00
- 44. verpleegster (in psych. inrichting), 1,00
- 42. telefoniste, 0,00
- 26. kantoorassistente, 0,00
- 7. controleuse in de industrie, 0,00

cluster 2 (R) K.R. (21) = 0,675

- 4. verbaal, 0,87
- 13. persoonlijk voorkomen, 0,87
- 45. verkoopster, 1,00
- 37. serveerster, 1,00
- 1. assistent jeugdleidster, 1,00
- 2. assistent groepleidster, 1,00
- 11. dierenarts-assistente, 1,00
- 12. dokters-assistente, 1,00
- 8. kassière, 1,00
- 9. konduktrice, 1,00

cluster 3 (R) K.R. (21) = 0,85

- 6.. kunstzinnig, 0,88
- 18. vormgeving belangrijker dan handvaardigheid, 0,88
- 8. vormgevoel, 0,87
- 5. bloemschikster, 1,00
- 13. etaleuse, 1,00
- 18. hoedenmaakster, 1,00
- 20. kostuummaaister, 1,00
- 25. kapster, 1,00

Specifieke clusters werden gevormd door

- 7. materiaalgevoel

- 15. handvaardigheid belangrijker dan toewijding
- 3. organisatorisch
- 17. handvaardigheid belangrijker dan aandacht
- 2. technisch
- 12. handvaardigheid

Bij vergelijking met de Q-analyse blijkt cluster 1 (R) een getransponeerde, bipolaire faktor van 3 (Q) en 4 (Q) te zijn, cluster 2 (R) overeen te stemmen met cluster 2 (Q) en cluster 3 (R) met 1 (Q).

De R-analyse geeft dus 4 van de 8 clusters van de Q-analyse. De informatie is dus in dit geval van R duidelijk minder. De reden hiervan is gelegen in de scheefheid van de datamatrix. Met 47 elementen kunnen makkelijker meer niet-specifieke clusters worden gevormd dan met 18 operatoren.

Bevestigd wordt dus de verwachting, dat R- en Q-analyse van de data geen grote verschillen zullen geven. De informatie, die de clusters geven, is echter wel anders. Terwijl de R-analyse inzicht geeft over de wijze waarop de beoordelaars te werk zijn gegaan, geeft de Q-analyse informatie over de te onderscheidde groepen van beroepen. Dit komt onder andere tot uitdrukking bij vergelijking van cluster 1 (R) met de clusters 3 (Q) en 4 (Q). De beoordelaars gaan er blijkbaar vanuit, dat wanneer 'fysieke inspanning' en 'toewijding' relevante beroepseisen zijn 'nauwkeurigheid' en 'aandacht' als eisen uitgesloten zijn en omgekeerd. Bij analyse van de beroepen blijkt deze 'wetmatigheid' voor de onderzochte beroepen niet op te gaan. Er is sprake van een groep verzorgende beroepen, waarvoor 'toewijding' en 'fysieke inspanning' belangrijke eisen zijn en een daarvan onafhankelijke groep beroepen waar 'nauwkeurigheid' en 'aandacht' als eisen centraal staan, gepresenteerd door respectievelijk cluster 4 (Q) en cluster 3 (Q).

5.2.3. Interindividueelonderzoek

Bij dit soort van onderzoeken gaat het meestal om *constructies van tests of vragenlijsten*. Dit is een traditioneel toepassingsgebied voor de clusteranalyse. Motivatie en satisfactie onderzoek vallen hier ook onder (zie Boon van Ostade, 1968). De analyse is hiervoor door Pennings (1966) en Hermans (1967) gebruikt. De eerste om te komen tot een vragenlijst voor 'satisfactie', 'intrinsieke-extrinsieke werkmotivatie', 'houding collega's' en 'promotiebehoefte'. De tweede voor de constructie van een vragenlijst voor 'prestatie-motivatie', 'faalangst' en 'sociale wenselijkheid' (zie ook Nederlands Instituut van Psychologen resp. 17.7., p. 180, 18.6., 18.21. en 23.7. p. 207, p. 216 en p. 259).

Voor voorbeelden wordt hier verder naar bovengenoemde literatuur verwezen, en naar de kapittel-enquête (zie paragraaf 4.1.) en het A.M. onderzoek (zie paragraaf 4.5.).

5.2.4. Type-onderzoek

Toepassing van de clusteranalyse voor onderzoek van types komt vooral bij data van *projectie-tests of observatie-tests* in aanmerking. Klassificatie van de protocollen op grond van overeenkomstige patronen van antwoorden maakt het mogelijk te komen tot een typen-indeling van de individuen. Ook bij *enquêtes* kan deze onderzoeksvorm worden toegepast. Aan het eind van deze paragraaf zal daar verder op worden ingegaan.

Een voorbeeld van een dergelijk onderzoek is gepubliceerd door Kempen en Boon van Ostade (1968). Hier zal meer op de methodologische zijde van dit onderzoek worden ingegaan en verder niet of nauwelijks op de inhoudelijke kant.

Het gaat daarbij om door Lutte en Sarti (1969), Lutte, Mönks, Kempen en Sarti (1969) verzamelde opstellen bij jongeren van 10 t/m 17 jaar in zeven Europese landen. De opdracht was een opstel te schrijven over de 'persoon op wie ik het liefst zou willen lijken' (Havighurst and MacDonald, 1959). De opstellen waren gecodeerd met nul (afwezig) en één (aanwezig) van 67 kenmerken, zoals de beschrijving van een 'man', van een 'vrouw', van een 'heldhaftige figuur', van de 'kleding', van het 'voorkomen', van de 'karaktereigenschappen'. Om te komen tot een empirische klassificatie werd bij 181 opstellen een Q-analyse uitgevoerd. De datamatrix was dus 181 bij 67 (zie bijlage II.4.). Deze klassificatie is meer in overeenstemming met het doel van het onderzoek: het ego-ideaal (de persoon op wie ik het liefst zou willen lijken). Het opstel wordt in dit onderzoek niet als prestatietest beschouwd. Om die reden is er ook geen bezwaar om korte en niets-zeggende opstellen uit het materiaal te verwijderen. Hierdoor is het mogelijk de variantie van de opstellen zoveel mogelijk te standaardiseren; $M = 0,12$, $S.D. = 0,04$, range 0,22 tot en met 0,06. De variantie van de

kenmerken is: $M = 0,11$, $S.D. = 0,06$, range 0,23 tot en met 0,01.

Ook de informatie is goed; 181 over 67. Een probleem vormt weer de statistische betrouwbaarheid. Tabel 4.6.1. leert dat een overeenkomstindex van 0,40 nodig is om de toetsing op generaliseerbare resultaten te kunnen uitvoeren. Het onderzoek is ook een voorbeeld van de toepassing van een hiërarchisch-onderzoek (zie paragraaf 5.3.2.). Er zijn namelijk op dezelfde data twee analyses uitgevoerd: één met een overeenkomstindex van 0,30 (Analyse A) en één met een index van 0,40 (Analyse B). De tweede analyse maakt het mogelijk de toetsing op toevalsresultaten uit te voeren. Deze is als volgt: $P\{r \geq 0,40\} = 0,00088$. Bij toeval kunnen dus bij deze grens 14 'significante' correlaties worden gevonden. 137 werden er gevonden. De nulhypothese kan dus worden verworpen ($t = 32,4$). Er is geen sprake van een artificieel resultaat. Ook inspectie van de gevonden clusters op hun psychometrische betrouwbaarheden geeft geen reden te vermoeden dat bepaalde onderdelen kunstmatig zijn, met uitzondering van cluster 13 uit analyse B. In tabel 5.2.2.1. staan de gevonden clusters met hun K.R.(21) uitkomst.

Wanneer met behulp van bijlage I.4. de betrouwbaarheidsgrens van de K.R.(21) wordt berekend dan blijkt met een N van 67, de χ^2 98,5 te zijn voor de één procentsdrempel. De grens wordt dan $n/n-1 \cdot 0,32$ en dus bij een cluster van twee operatoren 0,64 ($n=2$) en van drie 0,48. Alleen cluster 13 van analyse B heeft een K.R.(21) waarde, die hieronder ligt. De K.R.(21) van dit cluster is echter wel significant voor 5%. De betrouwbaarheidsgrens is dan 0,52.

Tabel 5.2.2.1 clusters

Analyse A

| nummer | aantal opstellen | K.R. (21) | nummer | aantal opstellen | K.R. (21) |
|--------|---------------------|-----------|--------|---------------------|-----------|
| 1 | 96 | 0,957 | 8 | 1 | -- |
| 2 | 53 | 0,914 | 9 | 3 | 0,628 |
| 3 | 9 | 0,749 | 10 | 1 | -- |
| 4 | 6 | 0,674 | 11 | 1 | -- |
| 5 | 4 | 0,715 | 12 | 1 | -- |
| 6 | 1 | -- | 13 | 2 | 0,724 |
| 7 | 1 | -- | 14 | 1 | -- |

Analyse B

| no. | aantal | K.R. (21) | no. | aantal | K.R. (21) | no. | aantal | K.R. (21) |
|-----|--------|-----------|-----|--------|-----------|-----|--------|-----------|
| 1 | 49 | 0,950 | 16 | 2 | 0,659 | 31 | 1 | -- |
| 2 | 25 | 0,904 | 17 | 2 | 0,657 | 32 | 1 | -- |
| 3 | 25 | 0,914 | 18 | 3 | 0,668 | 33 | 1 | -- |
| 4 | 5 | 0,706 | 19 | 4 | 0,714 | 34 | 1 | -- |
| 5 | 1 | -- | 20 | 1 | -- | 35 | 1 | -- |
| 6 | 3 | 0,726 | 21 | 1 | -- | 36 | 1 | -- |
| 7 | 6 | 0,725 | 22 | 1 | -- | 37 | 4 | 0,814 |
| 8 | 4 | 0,698 | 23 | 1 | -- | 38 | 1 | -- |
| 9 | 2 | 0,652 | 24 | 1 | -- | 39 | 1 | -- |
| 10 | 1 | -- | 25 | 1 | -- | 40 | 1 | -- |
| 11 | 11 | 0,847 | 26 | 1 | -- | 41 | 1 | -- |
| 12 | 1 | -- | 27 | 1 | -- | 42 | 1 | -- |
| 13 | 2 | 0,594 | 28 | 3 | 0,725 | 43 | 1 | -- |
| 14 | 4 | 0,742 | 29 | 1 | -- | 44 | 1 | -- |
| 15 | 1 | -- | 30 | 1 | -- | | | |

Alleen de eerste twee clusters van analyse A en de drie eerste clusters van analyse B zijn verder uitgewerkt. In resp. tabel 5.2.4.1. en 5.2.4.2. staan de kenmerken van deze groepen van opstellen. Vermeld zijn alleen die kenmerken, waarvan de klassificatiemaat discriminerend is, dat wil zeggen waarvan de maat significant met minstens één ander cluster verschilt ($p < 0,05$ bij een t-toets voor proporties). De klassificatiemaat is gegeven in percentage van de frequentie van voorkomen van het kenmerk.

Op grond van de discriminerende kenmerken zijn de clusters omschreven als cluster 1 (A): mannelijke ideaaltype

cluster 2 (A): vrouwelijke ideaaltype

cluster 1 (B): helden

cluster 2 (B): zorgzame moederfiguren

cluster 3 (B): voorbeeldige leeftijdsgenoten

cluster 4 (B): beroemdheden van het lichte genre.

Bij een enquête gaat het meestal om beschrijvingen van een situatie door personen. Het is gebruikelijk dit via geprecodeerde vragen en schalen te doen en de antwoorden van de beoordelaars op één bepaalde vraag samen te nemen in de vorm van percentages en gemiddelden en wel als beschrijving van de 'situatie'.

Door nu de clusteranalyse toe te passen als Q-analyse en dus als onderzoek van types wordt het mogelijk na te gaan of er homogene groepen van protocollen of beoordelaars zijn met specifieke patronen van antwoorden. Een enquête protocol kan namelijk worden opgevat als een individuele beschrijving van een stimulus; de situatie, zoals beoordeeld door deze persoon. Terwijl de vragen en schalen de operatoren of kenmerken zijn van deze stimulus, die als 'situatie' constant wordt gehouden. Deze constantheid is nu alleen bij homogene groepen van protocollen gewaarborgd en ook dan alleen is het verantwoord om deze 'situatie' door middel van percentages (klassificatiematen) te beschrijven.

Tabel 5.2.4.1.

Kenmerken die discrimineren tussen de twee hoofdclusters van
clusteranalyse A

| Kenmerk | 100 x klassificatiemaat | |
|------------------------------------------------|--------------------------|--------------------------|
| | Cluster 1 (N=96, 53%) | Cluster 2 (N=53, 29%) |
| Geslacht | | |
| 1. mannelijk..... | 100,0% | 0,0% |
| 2. vrouwelijk..... | 0,0% | 98,1 |
| Beroep | | |
| 16. sociaal..... | 26,0 | 50,9 |
| Vrije tijd | | |
| 24. lezen..... | 11,4 | 22,6 |
| 25. cultuur..... | 22,8 | 32,1 |
| 27. huishoudelijke activiteiten | 0,0 | 11,5 |
| 30. doorgebracht binnen organi- saties..... | 1,0 | 9,4 |
| Uiterlijk | | |
| 35. krachtig..... | 19,8 | 1,9 |
| 36. sportief..... | 2,1 | 9,4 |
| 37. groot van gestalte..... | 13,5 | 30,2 |
| 40. normale omvang..... | 0,0 | 15,1 |
| 43. donker haar..... | 9,4 | 22,6 |
| 48. gezichtsdetails..... | 12,5 | 24,5 |
| 49. kleding..... | 2,1 | 9,4 |
| Karakter | | |
| 52. fijngevoeligheid..... | 25,0 | 52,8 |
| 54. altruïsme..... | 43,7 | 24,5 |
| 58. doorzettingsvermogen..... | 16,7 | 3,8 |
| 61. moed..... | 9,4 | 0,0 |
| Waarden | | |
| 65. godsdienst..... | 33,3 | 17,0 |
| 67. roem, succes..... | 20,8 | 1,9 |

Tabel 5.2.4.2.

Kenmerken die discrimineren tussen de vier hoofdclusters van
clusteranalyse B

| Kenmerk | 100 x klassificatiemaat | | | |
|---------------------------|-----------------------------|-----------------------------|-----------------------------|----------------------------|
| | Cluster 1 (N=49, 27%) | Cluster 2 (N=25, 14%) | Cluster 3 (N=25, 14%) | Cluster 4 (N=11, 7%) |
| Geslacht | | | | |
| 1. mannelijk..... | 100,0% | 0,0% | 72,0% | 90,9% |
| 2. vrouwelijk..... | 0,0% | 96,0 | 24,0 | 9,1 |
| Leeftijd | | | | |
| 13. jong..... | 10,2 | 12,0 | 96,0 | 72,7 |
| 14. volwassen..... | 77,5 | 80,0 | 0,0 | 0,0 |
| Beroep | | | | |
| 16. sociaal..... | 32,6 | 72,0 | 4,0 | 0,0 |
| 18. artistiek..... | 0,0 | 0,0 | 4,0 | 90,9 |
| 19. studerend..... | 2,0 | 0,0 | 68,0 | 0,0 |
| Vrije tijd | | | | |
| 21. sport..... | 6,1 | 0,0 | 48,0 | 45,4 |
| Uiterlijk | | | | |
| 34. knap, elegant..... | 20,4 | 24,0 | 40,0 | 36,4 |
| 35. krachtig..... | 26,5 | 0,0 | 8,0 | 0,0 |
| 43. donker haar..... | 4,1 | 24,0 | 24,0 | 9,1 |
| 44. haar, overige details | 4,1 | 8,0 | 28,0 | 9,1 |
| 48. gezichtsdetails..... | 4,1 | 20,0 | 24,0 | 9,1 |
| 49. kleding..... | 2,0 | 8,0 | 4,0 | 27,3 |
| Karakter | | | | |
| 52. fijngevoeligheid..... | 26,5 | 44,0 | 40,0 | 9,1 |
| 54. altruïsme..... | 61,2 | 28,0 | 32,0 | 0,0 |
| 56. plichtsbesef..... | 16,3 | 24,0 | 68,0 | 0,0 |
| 58. doorzettingsvermogen. | 18,4 | 0,0 | 0,0 | 0,0 |
| 61. moed..... | 18,4 | 0,0 | 0,0 | 0,0 |
| Waarden | | | | |
| 65. godsdienst..... | 40,1 | 20,0 | 36,0 | 0,0 |
| 67. roem,succes..... | 26,5 | 0,0 | 8,0 | 36,4 |

5.3. Vormen van toepassing

Het is minder gebruikelijk om klassificatiemethoden op te vatten als bruikbaar voor *conditie-onderzoek*. De reden hiervan is gelegen in het betrekkelijk eenvoudige model van faktoranalyse, dat ervan uitgaat dat een variabele afhankelijk gesteld kan worden van de som van een aantal onafhankelijke variabelen. Het model is wel multiconditioneel, echter linear en meestal onafhankelijk. Eventuele interacties tussen de factoren zoals bijvoorbeeld moderator-effecten zijn niet te bepalen. Een moderator-effect is een interactie tussen twee variabelen (Saunders 1954). Daarom wordt veelal de voorkeur gegeven aan variantieanalyse. Deze methode is echter alleen in experimenteel onderzoek te gebruiken, omdat alleen daar meestal met de onafhankelijke variabelen kan worden gemanipuleerd. Bij niet-experimentele research of veldonderzoek stuit deze methode op bezwaren. Lege cellen komen bij dergelijke onderzoeken vrij veel voor. Correlatie-onderzoek kan dan alleen worden toegepast (Hyman, 1964).

Onderzoek naar conditionele relaties wordt bijna altijd toetsend uitgevoerd door na te gaan in welke mate de data corresponderen met een aantal theoretisch gestelde verklaringsmogelijkheden. De inductieve methode is meestal niet toepasbaar vanwege het grote aantal verklaringen, die voor dezelfde correlatiestructuur kan worden gegeven. Blalock (1964) zegt hierover: "In order to avoid empiricist objections to causal terminology, we prefer to think in terms of causal models of reality..... Since these models do not refer to reality itself, and since a number of alternative models may yield the same predictions, we can never actually establish a given model. But we can proceed by eliminating or modifying inadequate models that give predictions inconsistent with the data" (p. 173 e.v.)

Bij constructie van schalen of tests wordt de iteratieve clusteranalyse in bovengenoemde zin gebruikt. Nagegaan wordt of de aanname juist is, dat de gepostuleerde eigenschap of het hypothetische construct een algemene of zogenaamde achterliggende conditie is. Wanneer dit zo is, dan moet bij een steekproef van elementen, die duidelijk verschillen in de gestelde eigenschap, er sprake zijn van overeenkomstige klassificaties door verschillende voor de eigenschap representatieve operatoren. Factoranalytisch wil dit zeggen, zoals al uiteengezet is paragraaf 3.5., dat er een algemene factor dient te zijn.

Ook *hiërarchisch-onderzoek* vraagt om een a priori opgestelde theorie. Voor *predictie-onderzoek* is dit echter niet nodig. Bij deze vorm van onderzoek dienen echter wel criteriumgegevens systematisch verzameld te zijn.

5.3.1. Conditie-onderzoek

De iteratieve clusteranalyse is bruikbaar voor de toetsing van de hypothese van een algemene factor in een set of subset van variabelen. Deze factor is de empirische manifestatie van een theoretisch gestelde conditie, die aan de uitslagen van de set of subset van variabelen ten grondslag ligt. In paragraaf 3.5. en 3.6. is aan deze toetsing al aandacht besteed.

De nul-hypothese is, dat geen enkele variabele met één van de andere variabelen correleert boven de gestelde overeenkomstindex. Hierbij dient deze index groter of gelijk te zijn aan de correlatie, die significant is op 5% niveau, en het resultaat van de analyse getoetst te worden op toeval, zoals uiteengezet in paragraaf 4.6. Door de herhaalde vergelijking van de intercorrelaties met de overeenkomstindex is het mogelijk ook bij toevalsdata één of meer correlaties te vinden, die boven de gestelde grens liggen.

Het is mogelijk, dat de nul-hypothese wordt verworpen, maar dat toch de hypothese van een algemene factor gefalsificeerd wordt en wel als het beste cluster relatief te klein is of te lage intercorrelaties heeft. In dat geval is de kans dat het cluster bij een volgende steekproef weer wordt gevonden te laag. Toetsing van de K.R.(21) of (20) (zie bijlage I.4., geval 1) geeft de mogelijkheid dit punt te evalueren.

Naast de toetsing op de convergerende validiteit is het ook nodig na te gaan of eventuele verkregen factoren ook de bedoelde condities representeren en dus de clusterscores niet correleren met andere niet-bedoelde factoren (condities). De divergerende validiteit van de set wordt dan getoetst.

Bij twee of meer toetsingen van de divergerende validiteit in één onderzoek zal met behulp van de in bijlage I.3. omschreven procedure allereerst moeten worden nagegaan of de correlaties, afgezien van hun richting, homogeen zijn. In tweede instantie kan de gemiddelde absolute intercorrelatie getoetst worden met de in geval 1 van bijlage I.1. vermelde procedure, die ook voor produktmoment correlaties geldt. Wij merken hierbij op, dat alleen wanneer de nul-hypothese wordt verworpen er een duidelijke uitspraak plaats vindt. Het niet-verwerpen van de nul-hypothese is een functie van de groote van de steekproef. Des te groter de steekproef wordt des te groter is de kans, dat de nul-hypothese wordt verworpen. Een adequatere toetsing is om van tevoren vast te stellen welke correlatie niet meer acceptabel is bijvoorbeeld de gebruikte overeenkomstindex en dan na te gaan of de correlaties lager zijn dan deze index. Verder dienen daarom altijd de absolute waarden te worden vermeld.

Het is mogelijk, dat op theoretische gronden een correlatie wordt verwacht met een andere conditie. Deze correlatie dient echter altijd lager te zijn dan de gestelde overeenkomstindex. Wanneer namelijk deze correlatie gelijk is aan of hoger dan de overeenkomstindex, is het onderscheid tussen beide operationaliseringën kunstmatig en onjuist. Zij vormen samen één cluster. De theorie wordt dus dan gefalsificeerd voor wat betreft de specificatie van de condities. Bij herhaalde falsificaties van de theorie moet aangenomen worden dat specificatie onmogelijk is en dus de theorie onjuist is. Wanneer de correlatie significant lager is en niet nul, wordt de theorie niet gefalsificeerd. Het is dan juist twee aparte condities te onderscheiden. Een dergelijk resultaat wijst echter wel op een mogelijke aanvulling van de theorie. Hiërarchisch gezien is het dan mogelijk beide condities op te vatten als onderdelen van een algemenere conditie.

Bovenstaande zal nader verduidelijkt worden met een vergelijking tussen de belangstellingsschalen van de BIT en de VIL voor eerste klas L.H.N.O.meisjes. De VIL is een voor deze groep geconstrueerde belangstellingstest (Scheerder en Schell, publicatie in voorbereiding). Hiervoor is de BIT afgenomen bij 128 meisjes en de VIL bij 1058 meisjes. Ter vergelijking is alleen de voorzijde van de BIT afgenomen, omdat dan de testtijd van de BIT en de VIL overeenkomstig zijn.

In tabel 5.3.1.1. staan de resultaten van de toetsing van de convergerende validiteit van de BIT-schalen. Opgeven zijn als pivots de items van de negen schalen, die een p/q verdeling hadden die niet extremer was dan 0,20/0,80 of omgekeerd. De L- en de M-schaal zijn niet bekeken, omdat deze schalen bestaan uit items van de andere schalen. De gebruikte overeenkomstindex is 0,30.

Tabel 5.3.1.1.
Toetsing convergerende validiteit
BIT bij 128 L.H.N.O.meisjes

| schaal | aantal opgegeven items | aantal items beste cluster | K.R.(20) |
|-----------------------|---------------------------|-------------------------------|----------|
| 1. sociaal | 9 | 9 | 0,82 |
| 2. agrarische arbeid | 9 | 9 | 0,80 |
| 3. literair-geestesw. | 9 | 9 | 0,67 |
| 4. amb. vormgeving | 9 | 5 | 0,54 |
| 5. administratief | 6 | 6 | 0,54 |
| 6. technisch-natuurk. | 6 | 5 | 0,51 |
| 7. handel | 8 | 4 | 0,45 |
| 8. voedselbereiding | 9 | 2 | 0,34 |
| 9. technische handv. | 5 | 1 | ---- |

De convergerende validiteit van 'technische handvaardigheid' en 'voedselbereiding' is duidelijk onvoldoende. In het laatste geval is de K.R.(20) niet significant (zie bijlage I.4., geval 1). Alleen van de eerste twee schalen kan gezegd worden dat de convergerende validiteit goed is. De andere vragen nog om verbetering.

Voor de VIL staan de uitkomsten in tabel 5.3.1.2. De analyse heeft op dezelfde wijze plaats gevonden als bij de BIT; toetsend met pivots van items met een p/q verdeling kleiner dan 0,20/0,80 en een overeenkomstindex van 0,30.

Tabel 5.3.1.2.
Toetsing convergerende validiteit
VIL bij 1058 L.H.N.O.meisjes

| schaal | aantal opgeven items | aantal items beste cluster | K.R. (20) |
|----------------------------------|----------------------------|----------------------------------|-----------|
| 1. representatief- verzorgend | 25 | 15 | 0,86 |
| 2. kunstzinnig | 16 | 16 | 0,86 |
| 3. agrarisch | 14 | 13 | 0,77 |
| 4. huishoudelijk | 15 | 7 | 0,71 |
| 5. vormgevend | 16 | 5 | 0,61 |
| 6. administratief | 13 | 5 | 0,56 |

De convergerende validiteit is voor de eerste vier schalen goed. De vijfde en zesde schaal behoeven echter nog verbetering.

Voor de toetsing van de divergerende validiteit zijn de intercorrelaties van de schalen van beide tests onderling berekend. In tabel 5.3.1.3. staat deze correlaties voor de BIT.

Tabel 5.3.1.3.
Intercorrelaties BIT-schalen
bij 128 L.H.N.O.meisjes

| schaal | 1 | 2 | 3 | 4 | 5 | 6 |
|--------|-------|-------|-------|-------|------|------|
| 1 | | | | | | |
| 2 | -0,31 | | | | | |
| 3 | -0,49 | -0,18 | | | | |
| 4 | -0,25 | -0,13 | 0,05 | | | |
| 5 | -0,31 | -0,10 | -0,23 | -0,20 | | |
| 6 | -0,22 | -0,08 | 0,17 | -0,25 | 0,25 | |
| 7 | -0,20 | -0,03 | -0,05 | -0,19 | 0,08 | 0,08 |

De divergerende validiteit van de BIT is onvoldoende. De gemiddelde absolute waarde is 0,18 en de gemiddelde waarde -0,12. Toetsing van de correlaties zegt in dit geval betrekkelijk weinig. Tengevolge van de kleine steekproef (128) is zelfs het grootste absolute verschil van r_{27} met de overeenkomstindex niet significant. Dit verschil heeft een t-waarde van 3,0 (zie eerste formule, geval 2a, bijlage I.1.). Voor de verwerping van de nul-hypothese is bij 21 herhalingen (het aantal intercorrelaties) volgens de formule van Ryan (1960, zie bijlage I.2.) een t-waarde van minstens 3,6 vereist bij een significantie drempel van 5%. Opvallend is het grote aantal negatieve correlaties. De BIT is een keuzetest. Om die reden kunnen lage negatieve intercorrelaties tussen de schalen worden verwacht. De intercorrelatie

tussen 'sociaal' en 'literair-geesteswetenschappelijk' is dan echter nog te hoog (0,49). Een betere presentatie is denkkelijk een bipolaire schaal, overeenkomstig 'representatief-verzorgend' van de VIL.

De intercorrelaties van de VIL staan in tabel 5.3.1.4.

Tabel 5.3.1.4.
Intercorrelaties VIL-schalen
bij 1058 L.H.N.O.meisjes

| schaal | 1 | 2 | 3 | 4 | 5 |
|--------|-------|------|------|------|------|
| 1 | | | | | |
| 2 | -0,10 | | | | |
| 3 | -0,22 | 0,15 | | | |
| 4 | -0,26 | 0,15 | 0,08 | | |
| 5 | -0,15 | 0,20 | 0,00 | 0,23 | |
| 6 | 0,07 | 0,10 | 0,07 | 0,13 | 0,05 |

De divergerende validiteit van deze test is beter. Alleen de correlaties van schaal 1 en 4, schaal 4 en 5, en schaal 1 en 3 zijn niet significant lager dan de overeenkomstindex. In tabel 5.3.1.5. staan de uitkomsten van deze toetsing. De gevolgde procedure staat beschreven in bijlage I.2.. De vergelijking is conform de eerste formule van geval 2a uit bijlage I.1. De gemiddelde absolute waarde is 0,13 en de gemiddelde waarde 0,03.

De conclusie is, dat de BIT zowel ten aanzien van de convergerende validiteit, zowel als de divergerende validiteit bij eerste klas L.H.N.O.meisjes minder goed voldoet dan de VIL.

Tabel 5.3.1.5.

Toetsing intercorrelaties VII
(zie tabel 5.3.1.4.)

| vergelijking | t | α | t_{α} | uitkomst |
|---------------------------------|-----|----------|--------------|-----------------------------------|
| $ r_{35} - \rho = 0,00-0,30 $ | 9,7 | 0,00041 | 3,3 | $t > t_{\alpha}$ significant |
| $ r_{56} - \rho = 0,05-0,30 $ | 8,1 | 0,00044 | 3,3 | $t > t_{\alpha}$ significant |
| $ r_{63} - \rho = 0,07-0,30 $ | 7,4 | 0,00048 | 3,3 | $t > t_{\alpha}$ significant |
| $ r_{34} - \rho = 0,08-0,30 $ | 7,1 | 0,00056 | 3,2 | $t > t_{\alpha}$ significant |
| $ r_{12} - \rho = 0,10-0,30 $ | 6,5 | 0,00062 | 3,2 | $t > t_{\alpha}$ significant |
| $ r_{26} - \rho = 0,10-0,30 $ | 6,5 | 0,00069 | 3,2 | $t > t_{\alpha}$ significant |
| $ r_{46} - \rho = 0,13-0,30 $ | 5,5 | 0,00080 | 3,2 | $t > t_{\alpha}$ significant |
| $ r_{23} - \rho = 0,15-0,30 $ | 4,9 | 0,00089 | 3,1 | $t > t_{\alpha}$ significant |
| $ r_{24} - \rho = 0,15-0,30 $ | 4,9 | 0,0010 | 3,1 | $t > t_{\alpha}$ significant |
| $ r_{15} - \rho = 0,15-0,30 $ | 4,9 | 0,0011 | 3,1 | $t > t_{\alpha}$ significant |
| $ r_{25} - \rho = 0,20-0,30 $ | 3,2 | 0,0015 | 3,0 | $t > t_{\alpha}$ significant |
| $ r_{13} - \rho = 0,22-0,30 $ | 2,6 | 0,0020 | 2,9 | $t < t_{\alpha}$ niet significant |

Opvallend is het wegvallen van cluster 63, 68 en 66 (B 6) bij analyse met 0,30. Bovenstaande schema wordt nauwkeuriger weergegeven als de respectievelijke clusters naast elkaar worden geplaatst. In tabel 5.2.2.3. is dit weergegeven.

Tabel 5.2.2.3.

Vergelijkbare clusters uit analyse A en B

| | A 1 | B 1 | B 25 | B 6 |
|----|-------------|-------------|-------------|-------------|
| 65 | 0,20 | 0,08 | 0,15 | 0,39 |
| 70 | <u>0,44</u> | <u>0,45</u> | 0,17 | 0,25 |
| 64 | <u>0,48</u> | <u>0,52</u> | 0,27 | 0,34 |
| 62 | <u>0,36</u> | 0,29 | <u>1,00</u> | 0,35 |
| 63 | <u>0,41</u> | 0,39 | 0,35 | <u>0,85</u> |
| 68 | 0,07 | -0,06 | 0,09 | <u>0,79</u> |
| 66 | 0,14 | 0,12 | 0,27 | <u>0,55</u> |
| 67 | 0,12 | 0,11 | 0,27 | 0,40 |
| 61 | 0,01 | -0,09 | -0,01 | 0,25 |
| 60 | 0,28 | 0,14 | 0,13 | 0,28 |
| 69 | 0,18 | 0,13 | 0,03 | 0,39 |
| | A 2 | B 6 | B 17 | B 34 |
| 65 | 0,03 | 0,39 | -0,02 | -0,02 |
| 70 | 0,23 | 0,25 | 0,02 | +0,19 |
| 64 | 0,01 | 0,34 | -0,06 | 0,20 |
| 62 | 0,05 | 0,35 | -0,06 | 0,27 |
| 63 | 0,07 | <u>0,85</u> | 0,03 | 0,27 |
| 68 | <u>0,40</u> | <u>0,79</u> | 0,39 | 0,42 |
| 66 | <u>0,45</u> | <u>0,55</u> | 0,01 | 0,36 |
| 67 | <u>0,39</u> | 0,40 | 0,29 | <u>1,00</u> |
| 61 | <u>0,49</u> | 0,25 | <u>0,87</u> | 0,30 |
| 60 | 0,11 | 0,28 | 0,23 | 0,22 |
| 69 | 0,06 | 0,39 | 0,23 | 0,12 |

| | A 3 | B 9 |
|----|-------------|-------------|
| 65 | 0,27 | 0,17 |
| 70 | 0,13 | 0,20 |
| 64 | 0,06 | 0,13 |
| 62 | 0,13 | 0,69 |
| 63 | 0,22 | 0,43 |
| 68 | 0,27 | 0,34 |
| 66 | 0,03 | 0,19 |
| 67 | 0,06 | 0,19 |
| 61 | 0,16 | 0,19 |
| 60 | <u>0,68</u> | <u>0,85</u> |
| 69 | <u>0,57</u> | <u>0,89</u> |

Hierbij blijkt opstel 63 een duidelijke tussenvariabele te zijn: correlaties met A 1 in B 1 en B 25 en met A 2 in B 6. Een verdere toetsing van B 6 gaat in de richting van A 2 en vermoedelijk ten koste van de lading van 63. Het 'ambivalente' karakter van 63 komt ook tot uitdrukking in B 9. Met de kern van A 3 heeft 63 nog een correlatie van 0,43. Bij A 3 wordt deze correlatie echter 0,22.

5.3.3. Predictie-onderzoek

Over de 'dramatische' waardeverminderingen van multiple regressieformules bij controle validatie, werd in vroegere publicaties ingegaan (Boon van Ostade en Hermans, 1967, 1968, zie ook Hazewinkel, 1968 a en b). In het licht van de in-deze-studie ontwikkelde theorie kan nu het in-die-publicaties gebruikte begrip 'psychometrische betrouwbaarheid' een adequatere inhoud worden gegeven. De mogelijkheid om een criterium te prediceren met behulp van een multiple regressieformule berust op de aanname, dat de condities, die de uitslagen van het criterium bepalen, van tevoren zijn te meten. Hierbij wordt verder aangenomen, dat het criterium multiconditioneel wordt bepaald. Wanneer nu de relevante condities onvoldoende betrouwbaar (en valide) worden gemeten, wordt de voorspellingsmogelijkheid van het criterium verkleind, hetgeen onder andere bij controle validatie blijkt. Hetzelfde gebeurt wanneer de condities niet constant werkzaam zijn in het criterium. De conclusie is, dat waardevermindering van multiple regressieformules voor psychologische data gezocht moet worden in de betrouwbaarheid of constantie van het criterium en in de betrouwbaarheid of validiteit van de predictoren. Een benadering van beide genoemde betrouwbaarheden kan worden verkregen door middel van de psychometrische betrouwbaarheid en wel als 'test-retest' bij het criterium en 'parallel-test' bij de predictoren. Ten aanzien van de psychometrische betrouwbaarheid van predictoren kan verder worden verwezen naar paragraaf 3.6.

Bij gebruik van de iteratieve clusteranalyse voor predictie-onderzoek kunnen dus de criteria en predictoren het best samen exploratief worden geanalyseerd (zie Boon van Ostade, Hermans 1968). Wanneer het echter om één criterium gaat, kan dit criterium beter eerst als pivot

worden opgegeven en daarna alle predictoren, die boven een bepaalde grens met het criterium correleren, exploratief worden geanalyseerd. Hierdoor wordt het mogelijk het criterium te voorspellen met clusters die redelijk voldoen ten aanzien van de convergerende- en divergerende validiteit. Tevens kan dan inzicht worden verkregen in de aard van de relevante (en niet-relevante) condities.

Als voorbeeld zal hier kort een oriënterend onderzoek worden behandeld (Gastkemper, nog niet gepubliceerd). 66 vertegenwoordigers hadden drie in het nederlands vertaalde amerikaanse vragenlijsten ingevuld, namelijk de Sales Motivation Inventory, de Sales Comprehension Test en de Sales Attitude Check List (zie Buros, 1961, resp. 2064, 2061, Taylor, Schneider en Symons, 1953). Van deze vertegenwoordigers was het algemene verkoopcijfer over één jaar bekend. De items zijn nu bekeken op hun relatie met dit criterium, dat gecorrigeerd is voor niet-voorziene marktfluctuaties, en wel om na te gaan in hoeverre met deze items dit criterium is te prediceren.

Het aantal items van de drie tests bedraagt totaal 136. Ieder item heeft vier antwoordmogelijkheden. Omdat het verwachte commerciële antwoord niet bekend is (behalve uit de amerikaanse sleutels), is ieder antwoord-alternatief beschouwd als een operator. Het aantal variabelen is dus 544. Bij de opgave van de pivot moet dan rekening worden gehouden met het feit, dat de vier operatoren van één item niet-experimenteel onafhankelijk zijn en dus niet samen in één pivot mogen voorkomen. Bij controle op de p/q verdeling met als grens 0,20/0,80 vielen 321 operatoren uit. Het onderzoek gaat dus over 223 variabelen. Samen met het criterium en twee anamnese gegevens (namelijk leeftijd en vooropleiding) wordt dit totaal 226.

Bij opgave van het criterium als pivot bleken slechts 17 operatoren meer dan 0,20 te correleren met het criterium. Dit aantal is niet significant verschillend van wat toevallig aan dergelijke lage correlaties, bij zoveel predictoren en een zodanig kleine steekproef kan worden gevonden. Desondanks zijn 12 van de 17 operatoren als pivot explorerend geanalyseerd. De overige 5 konden niet in de pivot worden opgenomen, omdat het antwoord-alternatieven van al vertegenwoordigde items betrof.

Het resultaat van deze analyse staat in tabel 5.3.3.1. De overeenkomstindex is 0,20.

Tabel 5.3.3.1.
Resultaten explorerende analyse
pivot 12 operatoren

| no. | aantal items cluster | K.R. (21) | correlatie met criterium |
|-----|----------------------------|-----------|-----------------------------|
| 1 | 2 | 0,21 | 0,28 |
| 2 | 5 | 0,59 | 0,44 |
| 3 | 2 | 0,10 | 0,39 |
| 4 | 2 | 0,16 | 0,34 |
| 5 | 1 | — | 0,24 |

Alleen cluster 2 is betrouwbaar genoeg om eventueel te gebruiken voor predictie. Wanneer de operatoren van dit cluster worden uitgeschreven (zie tabel 5.3.3.2.) dan is de indruk dat het hier om een behoefte aan 'initiatief' gaat. De vertegenwoordigers met een behoefte om structuur aan te brengen in de situatie zouden dus hogere verkoopcijfers hebben dan vertegenwoordigers die meer geneigd zijn zich te laten bepalen door hun omgeving.

Tabel 5.3.3.2.
Operatoren van cluster 2

| | tekst | lading (correlatie) |
|----|----------------------------------------------------------------------|---------------------|
| 1. | niet: omroeper zijn van internationale nieuwsberichten | 0,74 |
| | wel : leiding geven aan de technische dienst van een radio-omroep | (0,41) |
| | wel : teksten schrijven voor een radio- commentator | (0,39) |
| 2. | niet: toneelmeester zijn | 0,71 |
| | wel : lichttechnicus zijn | (0,25) |
| 3. | wel : ik kan goed leiding geven | 0,64 |
| 4. | wel : werk doen waarbij je veel verantwoordelijkheid hebt | 0,63 |
| | niet: een baan hebben die je zekerheid biedt | (0,27) |
| 5. | niet: ik schrijf mijn rapporten niet op tijd | 0,41 |
| | wel : ik weet meer dan voor mijn werk noodzakelijk is | (0,24) |

Wij merken hierbij op, dat bovengenoemd onderzoek een voorbeeld is van een vooronderzoek in de geest van De Groot (1961, p.324). Wanneer er geen verder onderzoek opvolgt, is alles van weinig waarde (zie ook eind van paragraaf 3.5.).

6. Vergelijking met de hoofdassenanalyse

6.1. Inleiding

Clustermethoden worden meestal opgevat als 'slappere' procedures dan factoranalytische. Lorr en Radhakrishnan (1967, p.47) beschouwen de methoden van McQuitty (1957, 1960 a, b en c, 1964, 1965, 1966 a en b), van Thorndike (1953), van Sawrey, Keller en Conger (1960), van Gengerelli (1963) en van Lorr, Klett en McNair (1963) als zodanig. Deze methoden zouden onnauwkeuriger en oppervlakkiger zijn dan de centroid analyse (Thurstone, 1947), hoofdassenanalyse (Harman, 1960) en de maximum likelihood van Lawley (1940, 1941).

Een tussenpositie nemen, volgens de literatuur, Tryon's cumulative communality clusteranalyse (1958) en de iteratieve factoranalyse van Wherry en Gaylord (1943) in. Hiermee is nog lang niet een uitputtend overzicht gegeven van bestaande cluster- en factormethoden. Zo is niet genoemd Bridges hiërarchische clusteranalyse (1966) en Elshout's (1963) methode van dezelfde naam (zie voor een overzicht van clustermethoden o.a. Bijnen 1969).

In deze studie is de iteratieve clusteranalyse niet vergeleken met alle bovenstaande methodieken. Alleen met de hoofdassenanalyse, ook wel voornaamste componenten analyse genoemd, is een vergelijking uitgevoerd. Deze beperking is om twee redenen ingevoerd. De eerste is dat nog steeds de hoofdassenanalyse (met varimax rotatie van factoren met eigenwaarden groter dan één en communaliteiten van één) de meest gebruikte factoranalytische techniek is. In paragraaf 6.2. wordt hier verder op ingegaan. De tweede reden is, dat het niet in de eerste plaats het doel van deze studie is om aan te tonen, dat nieuwe of andere methodieken dan de iteratieve clusteranalyse in bepaalde gevallen betere oplossingen geven.

De iteratieve clusteranalyse is, zoals zal worden gedemonstreerd, een adequatere methode voor klassificatie dan de hoofdassenanalyse als

- a. één algemene klasse (factor, cluster) of een duidelijk dominerende klasse wordt verwacht,
- b. er grote onzekerheid bestaat ten aanzien van het aantal te verwachten klassen,
- c. het aantal operatoren (variabelen) relatief groot is (meer dan 75) of relatief groter dan het aantal elementen (proefpersonen),
- d. de aanname van een normale verdeling op interval niveau van de kwantificeringen van de operatoren duidelijk misplaatst is,
- e. een hiërarchische klassificatie op prijs wordt gesteld, en
- f. de varianties van de operatoren geen grote onderlinge verschillen geven.

De hoofdassenanalyse daarentegen is meer van toepassing als

- a. er een aantal ongeveer even sterke factoren wordt verwacht,
- b. er een redelijke zekerheid bestaat ten aanzien van het aantal te verwachten factoren,
- c. het aantal variabelen niet te groot is, in absolute zin minder dan 150 of in relatieve zin gelijk of minder dan het aantal proefpersonen,
- d. de aanname van kwantificeringen op interval niveau voor de variabelen acceptabel is, en
- e. er duidelijke verschillen in de varianties van de variabelen zijn en het niet zinvol of mogelijk is om deze varianties gelijk te maken.

De kritische punten van de hoofdassenanalyse zijn gelegen in

- a. de bepaling van het aantal factoren, en
- b. de resultaten bij bewerking van aselechte data.

Op beide punten zal nader worden ingegaan en aangetoond zal worden dat deze punten minder moeilijkheden geven bij gebruik van de clusteranalyse.

De kritische punten van de clusteranalyse worden gevormd door

- a. de bepaling van de ladingen, en
- b. het niet-geven van subclusters.

Daarnaast wordt in de literatuur kritiek gegeven op deze methode of soortgelijke methoden o.a. door Mosier (1936), Guilford (1954), Segers (1965, 1968), Tellegen (1968) en Bijnen (1969). Deze kritiek zal worden besproken en zo mogelijk worden weergelegd.

De vergelijking van de hoofdassenanalyse en clusteranalyse heeft plaats gevonden met behulp van zogenaamde simulatie-onderzoeken. Het gebruik van artificiele data ter vergelijking van methodieken wordt in de psychologie nog niet algemeen toegepast, zover ons bekend hebben Henrysson (1950), Lawley en Swanson (1954), Jöreskog (1963) en Browne (1968 b) deze methode toegepast. Meestal worden reële data gebruikt om methodieken te vergelijken. Het bezwaar van deze data is echter, dat de populatiewaarden onbekend zijn en daarmee een goed criterium ontbreekt. De onderzoeker is verder gebonden aan de steekproefgrootte, factorstructuur, het aantal variabelen van de datamatrix en de verdeling van de ware en geobserveerde scores. Met een betrekkelijk eenvoudig programma kan bij simulatie de onderzoeker al deze parameters variëren, terwijl precies bekend is wat de werkelijke waarden (factorladingen en factorscores) zijn. Deze voordelen wegen duidelijk op tegen het kunstmatig karakter van de analyse.

Het programma dat hier is gebruikt, is gemaakt door H.E.M. Borgers (I.B.M. 360-50, K.U., Nijmegen). Het gaat in 't kort als volgt te werk. De onderzoeker geeft een factorstructuur op bestaande uit m orthogonale factoren van n variabelen. Iedere variabele heeft in ieder van de m factoren een bepaalde lading. Teven wordt opgegeven de grootte van

de steekproef (N). Het programma trekt nu aselekt uit een normale verdeling met een gemiddelde van vijftig en een sigma van tien voor iedere 'proefpersoon' of element een factorscore voor ieder van de m factoren. Met behulp van deze scores en de lading wordt nu de ware score per variabele (operator) bepaald met behulp van de formule:

$$T_{bi} = a_{i1} \cdot F_{1b} + a_{i2} \cdot F_{2b} + \dots + a_{im} \cdot F_{mb}$$

Hierbij is T_{bi} de ware score van element b bij operator i,

$a_{i1}, a_{i2}, \dots, a_{im}$ de ladingen van operator i voor
de factoren 1, 2, ..., m, en

$F_{1b}, F_{2b}, \dots, F_{mb}$ de factorscores van element b voor
de factoren 1, 2, ..., m.

Deze ware score wordt nu omgezet in een geobserveerde score en wel met behulp van de volgende formule:

$$t_{bi} = T_{bi} + \delta_{bi} \cdot \sqrt{1-h_i^2}$$

Hierbij is t_{bi} de geobserveerde score,

T_{bi} de ware score,

δ_{bi} een aselekt getal getrokken uit een normale verdeling
($M=50, \sigma=10$), en

h_i^2 de communaliteit: $h_i^2 = \sum_{j=1}^m a_{ji}^2$.

Wij wijzen erop dat deze simulering volledig uitgaat van het factoranalytisch model zoals gedefinieerd door Harman (1960, p.11 e.v., zie ook formule 5 in paragraaf 3.4.) en dus optimale data (normaalverdeling factorscores, ware- en geobserveerde scores in voldoende aantal eenheden) voor factoranalyse geeft.

Hierbij dient te worden opgemerkt, dat het factoranalytische model van de hoofdassenanalyse (zie o.a. Harman, p.154 e.v.) afwijkt

van bovengenoemde simulatie. Bij de hoofdassenanalyse wordt de geobserveerde score opgevat als de som van produkten van factorlading en factorscore voor evenveel factoren als er variabelen zijn. De unieke meetfout is in dit model dus nul. Op grond daarvan alleen al zou de hoofdassenanalyse als een inadequate methode kunnen worden beschouwd. Meestal is de onderzoeker echter geïnteresseerd in duidelijk minder factoren dan het aantal variabelen; $m < n$. Op grond daarvan wordt het dan gerechtvaardigd geacht om de hoofdassenanalyse als extractiemethode voor de rotatie te gebruiken. De overige $(n-m)$ factoren worden dan beschouwd als onbelangrijke of storende factoren.

De gesimuleerde data, waarmee beide analyses worden vergeleken, dienen echter overeenkomstig te zijn met psychologische data. Dat is voor een factoranalytisch model, dat niet uitgaat van een van tevoren vastgesteld aantal factoren en daarnaast voor iedere variabele een unieke factor onderscheidt, duidelijk meer het geval dan voor het model van de hoofdassen.

De data van de acht simulatie-onderzoeken zijn systematisch gevarieerd in hun factorstructuur. De eerste twee zijn aselechte data. De derde bestaat uit één algemene factor. De vierde uit één algemene factor en vijf kleine groepsfactoren. De vijfde uit twee exact gelijke groepsfactoren. De zesde en zevende uit drie groepsfactoren en één subfactor en de laatste uit drie groepsfactoren en drie subfactoren. Onder subfactoren worden algemene- of groepsfactoren verstaan van variabelen die allemaal ook ladingen hebben in sterkere factoren. De vijf groepsfactoren van simulatie IV zijn in dit opzicht ook subfactoren. Hiermee zijn de belangrijkste

variatiën in factorstructuur gegeven, zoals deze zover ons bekend bij psychologische data voorkomen. Onderzoeken waar meer dan zes factoren bij 14 tot 27 variabelen worden geëxtraheerd, zijn in de psychologie betrekkelijk zeldzaam. Het aantal variabelen is verder om praktische redenen tot 50 beperkt. Het gebruikte hoofdassenanalyseprogramma kan alleen ten koste van veel computertijd meer dan 50 variabelen, met een stop bij 75, analyseren. De ladingen zijn nooit groter gekozen dan 0,80 en kleiner dan 0,10. Hogere waarden zijn ons inziens weinig realistisch, lagere waarden weinig relevant. De communaliteiten zijn verder nooit hoger dan 0,86. De meeste waarden liggen duidelijk lager. Alle populatie factorstructuren staan, als zij niet in de tekst worden vermeld, in bijlage II.1. De steekproefgrootte is verder voor alle onderzoeken op honderd gehouden. Alleen voor het eerste onderzoek is deze twintig en wel om een vergelijking te kunnen doorvoeren met de analyse van de beroepen uit paragraaf 5.2.1. en voor onderzoek IV zeshonderd, ter controle van de K.R.-formule (zie paragraaf 3.6.).

Op grond van deze steekproefgrootte moet al verwacht worden, dat zowel de hoofdassenanalyse als de iteratieve clusteranalyse nooit exact de opgegeven factorladingen (en dus de gebruikte factorscores) zullen teruggeven. Het is daarbij belangrijk om na te gaan of de gebruikte steekproef in voldoende mate representatief is voor de populatie van mogelijke steekproeven. Op zuiver toevallige redenen kan de steekproef een sterk afwijkend karakter hebben en dus de resultaten van de analyse weinig representatief zijn voor de gekozen factorstructuur.

Deze controle heeft plaats gehad door empirisch de standaarddeviatie uit te rekenen van de verschillen tussen de in de dataset verkregen correlaties en de werkelijke populatiecorrelaties. Per definitie is de populatiecorrelatie gelijk aan de som van de producten van de ladingen (zie Harman, 1960, formule 2.34). Voor iedere factorstructuur kunnen dus alle populatie-intercorrelaties worden berekend. Deze waarden zijn omgezet in Fisher z-waarden en afgetrokken van de z-waarden van de berekende correlaties. Van deze $\frac{1}{2}n(n-1)$ verschillen, waarbij n het aantal variabelen is, is de standaarddeviatie berekend. Deze standaarddeviaties van de acht simulatie-onderzoeken staan in de eerste kolom van tabel 6.1.1. Tevens staan daar in de tweede kolom de theoretisch verwachte standaarddeviaties, die berekend zijn met de volgende formule (zie o.a. Guilford, 1965, 8.15)

$$\sigma_z = \frac{1}{\sqrt{N-3}}$$

waarbij N de grootte van de steekproef is.

Tabel 6.1.1.

Berekende en theoretisch verwachte standaarddeviaties (100 x) van de verschillen in z-waarden tussen de geobserveerde- en populatiecorrelaties van de simulatie-onderzoeken

| simulatie onderzoek | 100 x berekende standaarddeviaties | 100 x theoretische standaarddeviaties |
|------------------------|---------------------------------------|------------------------------------------|
| I | 25,18 | 24,25 |
| II | 9,83 | 10,15 |
| III | 9,47 | 10,15 |
| IV | 3,49 | 4,09 |
| V | 10,52 | 10,15 |
| VI | 8,93 | 10,15 |
| VII | 10,31 | 10,15 |
| VIII | 8,78 | 10,15 |

Vergelijking van de berekende standaarddeviaties en de theoretische standaarddeviaties leert, dat de data van de acht simulatie-onderzoeken in voldoende mate representatief zijn. Het grootste verschil is 0,0137 (simulatie VIII). Het gemiddelde absolute verschil is 0,0071, waarmee ook is aangetoond dat de aanname van onafhankelijkheid, die ook voor bovengenoemde vergelijking geldt, praktisch geen ernstige problemen oplevert, zoals Jones en Fiske (1953) stellen (zie ook paragraaf 4.6.).

De simulatie-onderzoekingen zijn gebruikt om de gestelde kritische punten van de hoofdassenanalyse en iteratieve clusteranalyse op hun praktische relevantie te onderzoeken. Omdat de statistische steekproefverdelingen van zowel 'ladingen' als 'aantal factoren' onbekend zijn en de gevonden verschillen dus niet op significantie kunnen worden getoetst, hebben wij ons beperkt tot zeer duidelijke verschillen en/of verschillen die in meerdere onderzoeken voorkomen. De opgegeven factorstructuur heeft daarbij steeds gefungeerd als criterium. De analysemethoden dienen deze structuur weer in de data terug te vinden. Iedere afwijking is een verschil in bovengenoemde zin.

6.2. Kritische punten van de hoofdassemblanalyse

Op het ogenblik is het zo, dat in Nederland factoranalyse door velen als een dubieuze methode wordt beschouwd. 'Het is de Rorschach van de psychometrici' is één van de meest duidelijke uitspraken in deze.

Wat is de reden van dit onbehagen? De ervaringen met het gebruik van factoranalyse zijn van dien aard, dat men vaak afziet van verdere toepassingen en oplossingen gaat zoeken in ordinale methodieken of parameter-vrije analyses dan wel zich sterk beperkt in gebruik van statistische methoden. Ter rechtvaardiging van factoranalyse moet hier wel worden gesteld, dat bovenstaande ervaringen vaak het gevolg zijn van een onjuist begrip en kennis van de beperkingen en voorwaarden, die aan de toepassing van factoranalyse ten grondslag liggen.

Welke zijn deze beperkingen en voorwaarden? Algemeen wordt tegenwoordig ingezien, dat de bepaling van het aantal factoren één van de belangrijkste oorzaken is van de problematiek (Dirkzwager, 1966, p.35 e.v.). Dit blijkt ook uit de literatuur. Regelmatig verschijnen artikelen, waar bestaande methodieken voor bepaling van dit aantal worden vergeleken en/of nieuwe geïntroduceerd. Wij noemen: McNemar (1942), Sokial (1959), Horn (1965), Cattell (1966), Pennel en Ahumada (1964), Humphreys (1964), Horn (1965), Cattell (1966), Pennel (1968), Browne (1968a en b). Vrij recent heeft Browne (1968b) drie van deze methoden met behulp van simulatie-onderzoek vergeleken. Zijn conclusie is, dat geen van deze stopregels voldoet. In zijn dertig simulatie-onderzoeken met een steekproefgrootte van 100, geeft de likelihood ratio test 24 maal het goede aantal factoren, de toets van Jöreskog 21 maal en de toets van Kaiser (eigenwaarden groter dan 1) 20 maal.

De kans op een onjuist resultaat zou dus één op de drie tot één op de vier analyses zijn. De likelihood test geeft daarbij te weinig factoren (in alle 6 gevallen), de toets van Kaiser teveel (in alle 10 gevallen). Bij 10 simulatie-onderzoeken met een steekproefgrootte van 1500 voldoen de likelihood test en de toets van Kaiser volledig en de toets van Jöreskog volledig niet. Dit resultaat is echter van weinig praktische waarde. Indien klassificatie-onderzoekingen alleen bij dergelijk grote steekproeven zouden kunnen plaatsvinden, dan zou de toepassing van psychologie een bijzonder kostbare kwestie worden.

De 'biased' schattingen van het aantal factoren veroorzaken nu vooral problemen, als er teveel factoren worden geschat. In dat geval namelijk geeft een rotatie naar 'simple-structure' (varimax rotatie) een duidelijk onjuist beeld en wel vooral wanneer de eerste factoren sterk domineren. Tenopyr en Michael (1963, 1964) hebben hier al op gewezen voor een ~~algemene~~ factor. "It appears that when applied to data having a large general factor, the varimax method appears to remove too much variance from the general factor and to redistribute the variance in such a manner as to preclude simple-structure" (1963, p.594). Wanneer nu deze redistributie plaatsvindt naar niet-bestaande factoren wordt het resultaat een volledig zinloze factormatrix. Tenopyr en Michael hebben ook ongunstige bevindingen met de quartimax (1963) en de biquartimin (1964) rotatiemethoden. Hun conclusie is dat de grafische methode nog het meest is te prefereren. Dit zou echter een stap achteruit betekenen. Gezocht wordt naar objectieve rotatiemethoden. Zo stelt Harman (1960): "From the very beginning of the application of simple-structure principles, it was recognized that the procedure was more of an art than a science. In an endeavor to put the rotations on a more objective basis, the first improvements were directed toward eliminating graphical procedures"(p.290).

Algemeen gesteld rotatie is naar 'simple-structure' alleen verantwoord, als verwacht mag worden, dat de datamatrix bestaat uit een aantal elkaar in sterkte niet teveel ontlopende factoren. In andere gevallen moet gezocht worden naar andere extra regels, zoals uiteengezet in paragraaf 3.5. van deze studie. Dit punt wordt meestal verdedigd door erop te wijzen dat mathematisch de intercorrelatiematrix door een familie van factormatrices kan worden verklaard. Hierdoor ontstaat echter de onbepaaldheid van de analyse en wordt het mogelijk door middel van rotatie van een willekeurig aantal factoren ieder resultaat te verkrijgen wat de onderzoeker wenst. Horn (1967) wijst op dit subjectieve karakter van factoranalyse. In zijn studie roteert hij 47 'random' variabelen met een steekproefgrootte van 300 naar een schijnoplossing van 21 factoren. Hoewel er in dit geval genoegzaam aanwijzingen zijn voor een (ervaren) analist, dat het hier een vreemd resultaat betreft (lage ladingen, hoge intercorrelaties tussen de factoren), kunnen we ons geheel verenigen met zijn conclusie: "There was thus indication that if an investigator were willing to interpret relatively low loadings of there seemed to 'make sense' he needn't bother to gather actual data: random variables may be labeled arbitrarily and pushed into solutions that make quite 'good sense'" (p.280).

Factoranalyse van aselechte data kan blijkbaar 'zinvolle' resultaten opleveren. Bij een goede uitvoering van de iteratieve clusteranalyse is dit niet mogelijk, zoals zal worden aangetoond.

Voor de vergelijking met de iteratieve clusteranalyse wordt gebruik gemaakt van die factoranalytische techniek, die de grootste populariteit geniet, dat wil zeggen: hoofdassenanalyse met communaliteiten van één, de stop van Kaiser (eigenwaarde kleiner dan één) en

varimax rotatie. Glass en Taylor (1966) vonden in de periode van 1963 t/m 1965 in vier Amerikaanse tijdschriften over onderwijs-research 46 studies waar gebruik was gemaakt van factoranalyse. Hiervan gaven 10 geen informatie over de gebruikte techniek, 23 gebruikten communaliteiten van één en 16 de hoofdassenanalyse met de stop van Kaiser en varimax rotatie.

Dit onderzoek is in deze studie herhaald bij negen tijdschriften (vijf Amerikaanse en vier Europese) voor de jaren 1966, 1967 en 1968. In tabel 6.2.1. staan de cijfers van Glass en Taylor en dit onderzoek voor resp. de extractie- en rotatietechniek, de communaliteiten en het aantal factoren.

Het eerste dat bij de vergelijking opvalt is, dat Glass en Taylor in vier tijdschriften van totaal 11 jaargangen evenveel toepassingen vinden als in dit onderzoek bij negen tijdschriften van totaal 27 jaargangen. Dit kan het gevolg zijn van de gekozen tijdschriften, slechts één tijdschrift is overlappend, namelijk Educational and Psychological Measurement. Het is echter moeilijk aan te nemen, dat in tijdschriften als Psychological Reports, Multivariate Behavioral Research, Journal of Applied Psychology, Journal of Educational Research, Occupational Psychology, Scandinavian Journal of Psychology, Nederlands Tijdschrift voor Psychologie en Bulletin de Psychologie, minder toepassingen worden gepubliceerd dan in American Educational Research Journal, Journal Educational Measurement en Journal Educational Psychology. De algemene indruk (zie ook begin van deze paragraaf) dat de toepassing van factoranalyse afneemt, wordt hiermee bevestigd. Deze vermindering gaat voorlopig niet ten voordele van clusteranalyse. In bovengenoemde negen tijdschriften komen totaal 8 toepassingen van clusteranalyse voor.

Tabel 6.2.1.

Vergelijking met de bevindingen van
Glass en Taylor (1966) ten aanzien
van gebruikte f.a. technieken

| | extractietechniek | | | rotatietechniek | |
|-----------------|-------------------|---------|--------------|-----------------|---------|
| | '63-'65 | '66-'68 | | '63-'65 | '66-'68 |
| hoofdassen | 27 | 38 | varimax | 28 | 24 |
| centroïd | 9 | 6 | var.+andere | 0 | 8 |
| hoofd.+centr. | 0 | 1 | andere rot. | 8 | 14 |
| geen inf. | 10 | 1 | geen inf. | 10 | 0 |
| | communaliteiten | | | aantal factoren | |
| | '63-'65 | '66-'68 | | '63-'65 | '66-'68 |
| comm. 1 | 23 | 13 | t.v. Kaiser | 16 | 11 |
| geschatte comm. | 13 | 16 | andere toets | 20 | 13 |
| geen inf. | 10 | 17 | intuïtief | | 16 |
| | | | geen inf. | | 6 |

Aan tabel 6.2.1. is af te lezen, dat de hoofdassenanalyse en de varimax rotatie de meest populaire f.a.methoden zijn, al is er een tendens aanwezig om naast varimax ook andere rotatietechnieken te gebruiken. Toepassing van communaliteiten van één ligt echter duidelijk anders. Opvallend is het grote aantal publicaties dat ten aanzien van dit punt geen informatie verschaft. Ook de toets van Kaiser voor het aantal factoren is niet zo populair. Het aantal factoren wordt meestal, zo is de indruk, intuïtief vastgesteld.

Bovenstaande overzichten maken duidelijk dat bij de 'niet-praten-maar-doen' research nog steeds de hoofdassenanalyse met varimax rotatie het meest wordt toegepast. Meer verdeelt lijkt de houding te zijn geworden (of bij de auteurs van artikelen in de door ons onderzochte tijdschriften) ten opzichte van de communaliteiten en ten aanzien van het aantal factoren. Of de gekozen alternatieven echter te prefereren zijn is nog een open vraag.

Algemeen wordt aangenomen dat de methode van Lawley (1941) duidelijk beter is. Deze methode is echter erg tijdrovend en vaak niet ter beschikking. Van Browne (1968b) krijgt men de indruk dat de hoofdasßenanalyse met de likelihood ratio test redelijk ten aanzien van het probleem van het aantal factoren zou kunnen voldoen. De hierna volgende simulatie-onderzoekingen suggereren, dat de hoofdasßenanalyse goed tot zeer goede resultaten geeft bij grote steekproeven ($N > 500$) én bij overpresentatie van de factoren, wanneer het aantal variabelen, dat een lading groter dan 0,40 in één factor heeft, meer dan twintig bedraagt.

Het is echter niet de bedoeling van deze studie om eventuele betere f.a. methodieken toe te passen, zolang deze mogelijke verbeteringen niet systematisch zijn onderzocht. Voldoende is om aan te tonen dat de gebruikelijke f.a.techniek problemen geeft bij de bepaling van het aantal factoren en bij analyse van aselechte data, en dat deze moeilijkheden niet voorkomen bij gebruik van de clusteranalyse.

Gebruikt zijn voor de hoofdasßenanalyse de onderdelen 'principale componenten analyse' (optie 8) met énen op de diagonaal en 'varimax rotatie' (optie 14). In dat geval worden alleen factoren met eigenwaarde groter dan één geëxtraheerd en geroteerd. Bij sommige bewerkingen wordt ook het 'minimale en maximale aantal factoren' (optie 11) opgegeven. Bovengenoemde programma's zijn ontwikkeld en uitgegeven door E.E.Ch.I Roskam en A. Horsten (1968, I.B.M. 360-50, K.U., Nijmegen).

De iteratieve clusteranalyse is bij alle acht onderzoeken exploratief uitgevoerd met een overeenkomstindex van 0,20 of 0,30 (behalve simulatie I en II). Daarna werd bij een aantal onderzoeken bepaalde sets van variabelen controlerend getoetst.

6.2.1. Het aantal factoren

In tabel 6.2.1.1. staan de simulatie-onderzoeken: hun steekproefgrootte (N), het aantal variabelen (n), een omschrijving van de factorstructuur, het aantal factoren en subfactoren. Onder subfactoren worden factoren verstaan die bestaan uit variabelen welke alle ladingen hebben in voorafgaande sterkere factoren. De laatste twee kolommen geven het aantal factoren dat respectievelijk de hoofdasßenanalyse en de iteratieve clusteranalyse van de data heeft bepaald.

Tabel 6.2.1.1.

De simulatie-onderzoeken met het door de hoofdasßen- en iteratieve clusteranalyse aantal bepaalde factoren (laatste twee kolommen)

| | N | n | factorstructuur | fact. | sub.f. | hoofdassen | iteratieve |
|------|-----|----|---------------------------|-------|--------|------------|------------|
| I | 20 | 50 | aselecte data | 0 | 0 | 15 | 0 |
| II | 100 | 50 | aselecte data | 0 | 0 | 21 | 0 |
| III | 100 | 30 | één alg. factor | 1 | 0 | 10 | 1 |
| IV | 600 | 27 | één alg.f. en 5 sub.f. | 1 | 5 | 6 | 1 |
| V | 100 | 12 | twee gelijke groepsf. | 2 | 0 | 3 | 2 |
| VI | 100 | 14 | drie groepsf. en 1 sub.f. | 3 | 1 | 4 | 3 |
| VII | 100 | 14 | drie groepsf. en 1 sub.f. | 3 | 1 | 4 | 3 |
| VIII | 100 | 14 | drie groepsf. en 3 sub.f. | 3 | 3 | 4 | 3 |

Alleen bij simulatie IV, VI en VII geeft de toets van Kaiser het aantal factoren exact weer. Het eerstgenoemde onderzoek heeft daarnaast in verschil met de andere onderzoeken een steekproefgrootte van 600. Dit resultaat is duidelijk minder gunstig dan van Browne (1968b). Ook wordt niet bevestigd dat de toets van Kaiser altijd evenveel of teveel factoren

zou geven. Browne's onderzoek is echter gebaseerd op vier groepsfactoren ($N=100$, $n=12$, $n=16$); een factorstructuur die overeenkomt met simulatie VI en VII.

De clusteranalyse geeft in alle gevallen exact het aantal factoren. Deze analyse is niet in staat subfactoren te bepalen. Daarbij dient wel te worden opgemerkt dat specifieke clusters conform de bevinding in paragraaf 4.1. zijn weggelaten, evenals niet-significante clusters.

Een meer gedetailleerd inzicht zal hier worden gegeven met behulp van de resultaten van simulatie V. In tabel 6.2.1.2. staan de populatiefactoren van dit onderzoek, samen met de door de hoofdas- senanalyse geëxtraheerde (E) en geroteerde factoren (R) en de gevonden clusters door de iteratieve (I).

Tabel 6.2.1.2.

| Resultaten van simulatie V | | | | | | | | | | | | |
|----------------------------|-------------|-----------------|-------|--------|-------------|-----------------|-------|--------|-------------|-----------------|-------|-------|
| factor | | | | factor | | | | factor | | | | |
| | hoofd-assen | itera- tieve | | | hoofd-assen | itera- tieve | | | hoofd-assen | itera- tieve | | |
| | I | E | R | I | II | E | R | I | III | E | R | I |
| 1 | 0,60 | 0,70 | 0,73 | 0,57 | 0,10 | -0,22 | 0,26 | 0,06 | 0,00 | 0,25 | -0,02 | 0,36 |
| 2 | 0,50 | 0,63 | 0,70 | 0,60 | 0,00 | -0,34 | -0,06 | 0,15 | 0,00 | -0,13 | 0,21 | -0,02 |
| 3 | 0,40 | 0,64 | 0,50 | 0,56 | 0,20 | -0,04 | -0,03 | 0,13 | 0,00 | -0,49 | 0,64 | 0,18 |
| 4 | 0,50 | 0,58 | 0,72 | 0,58 | 0,00 | -0,39 | 0,10 | -0,05 | 0,00 | 0,27 | -0,16 | -0,04 |
| 5 | 0,60 | 0,64 | 0,75 | 0,69 | 0,10 | -0,37 | 0,07 | -0,01 | 0,00 | 0,15 | -0,02 | 0,10 |
| 6 | 0,80 | 0,81 | 0,78 | 0,72 | 0,00 | -0,25 | 0,05 | 0,18 | 0,00 | -0,18 | 0,36 | 0,12 |
| 7 | 0,10 | 0,44 | 0,06 | 0,04 | 0,60 | 0,52 | 0,49 | 0,62 | 0,00 | -0,12 | 0,48 | 0,22 |
| 8 | 0,00 | 0,28 | -0,14 | -0,03 | 0,50 | 0,56 | 0,25 | 0,55 | 0,00 | -0,50 | 0,75 | 0,12 |
| 9 | 0,20 | 0,39 | 0,10 | 0,19 | 0,40 | 0,40 | 0,43 | 0,17 | 0,00 | -0,04 | 0,35 | 1,00 |
| 10 | 0,00 | 0,36 | 0,08 | 0,05 | 0,50 | 0,42 | 0,54 | 0,54 | 0,00 | 0,16 | 0,16 | -0,02 |
| 11 | 0,10 | 0,31 | 0,02 | 0,13 | 0,60 | 0,52 | 0,76 | 0,67 | 0,00 | 0,45 | -0,05 | 0,08 |
| 12 | 0,00 | 0,39 | 0,08 | 0,18 | 0,80 | 0,51 | 0,71 | 0,62 | 0,00 | 0,31 | 0,09 | 0,10 |

De clusteranalyse is uitgevoerd met een overeenkomstindex van 0,20. Variabele 9 vormde een specifiek cluster (derde factor). Terwijl het derde cluster op grond van zijn specificiteit terzijde wordt gelegd, geeft de hoofdassenanalyse zowel bij de extractie als bij de rotatie een derde 'betekenisvolle' factor met ladingen van 0,40 of hoger bij drie variabelen.

De clusteranalyse schat de ladingen - in dit geval - ook nauwkeuriger. Het gemiddelde absolute verschil in z-waarden tussen populatie en berekende lading voor de eerste twee factoren, is bij de clusteranalyse 0,11 en bij de hoofdassenanalyse 0,15. Bij vari-max rotatie van alleen de eerste twee factoren wordt deze laatste waarde 0,13.

6.2.2. Aselecte data

De studie van Horn (1967) bracht naar voren dat f.a. van aselechte data 'zinvolle' resultaten zou kunnen geven. Gezien de wiskundig-statistische achtergrond van factoranalyse kan echter worden verwacht, dat de analysetechnieken hierop zijn gecontroleerd. Dit bleek echter niet het geval te zijn. Bij de twee door ons met behulp van de computer gegenereerde aselechte datasets van 50 bij 20 en 50 bij 100, respectievelijk van simulatie I en II, geeft de hoofdasenanalyse resp. 15 en 21 factoren! Gemiddeld hebben 4 (simulatie I) en 2 (simulatie II) variabelen ladingen van minstens 0,40 in deze factoren. Het resultaat is dus een groot aantal kleine groepsfactoren. Hieraan zou een aselechte datamatrix te herkennen zijn ware het niet dat dan geen onderscheid mogelijk is met een datamatrix die juist wel bestaat uit een verzameling van kleine groepsfactoren. Door controle validatie -herhaling van het onderzoek op een nieuwe steekproef- zou het misschien mogelijk zijn om te ontdekken dat de gevonden factoren artefacts zijn. Er is echter wel reden om hieraan te twijfelen gezien de vele rotatiemogelijkheden, waar de onderzoeker over beschikt en de afwezigheid van een toetsing voor de mate van overeenkomst tussen factormatrices.

Een adequater alternatief is volgens ons het gebruik van de clusteranalyse. In dat geval is de kans dat kunstmatige clusters psychologisch 'zinnig' worden geduid, veel kleiner. Toetsing van de K.R.(20) of (21) van de clusters geeft in eerste instantie de mogelijkheid de resultaten te controleren (zie bijlage I.4.). De tweede controle wordt gegeven door de toetsing op toeval, zoals beschreven in paragraaf 4.6.

Bovenstaande is nader te adstrueren met de resultaten van de clusteranalyse voor simulatie I en II. De eerste dataset bestaat uit 50 'operatoren' die ieder 20 'elementen' hebben geklassificeerd en wel geheel volgens toeval. Een dergelijke matrix mag door de hoofdassenanalyse wiskundig gezien niet worden geanalyseerd: het aantal 'proefpersonen' is kleiner dan het aantal 'variabelen'. De data zijn dan lineair afhankelijk, de determinant is nul. Deze beperking is uiteraard psychologisch niet verdedigbaar. De beperking is het gevolg van de matrix-algebra, die door Hotelling (1933) is geïntroduceerd om het probleem van de factorextractie op te lossen. De introductie van de matrix-algebra voor een betrekkelijk eenvoudige kwestie heeft, ons inziens, de werkelijke klassificatieproblematiek in de psychologie op een overbodige wijze gecompliceerd. De clusteranalyse, die niet uitgaat van matrix-algebra, kent deze beperking niet. Dataset's, waar het aantal elementen duidelijk kleiner is dan het aantal operatoren, worden echter statistisch gezien wel bijzonder gevoelig voor kunstmatige resultaten. Het is daarom belangrijk bij steekproeven, die kleiner zijn dan veertig, de resultaten op toeval te toetsen.

Simulatie I is ook gekozen om een vergelijking mogelijk te maken met de analyse van de beroepen (zie paragraaf 5.2.1.). In dat geval was de matrix 18 bij 47. Met behulp van o.a. tabel 4.1.1. werd een overeenkomstindex van 0,61 gekozen. De analyse gaf 4 niet-specifieke clusters. Hiervan bestonden 3 clusters uit 2 operatoren en 1 uit 3. Het aantal 'significante' correlaties is dus vijf. De uitkomsten van de K.R.(21) zijn voor alle clusters significant; voor 2 clusters voor de 1% sign. drempel. Bij deze controle vallen de resultaten dus niet door de mand. Dit is o.a. het gevolg van de hoge overeenkomstindex. Bij toetsing

op toeval blijkt echter het aantal 'significante' correlaties gelijk aan het gemiddelde aantal dat volgens toeval mag worden verwacht. De student's t van een correlatie van 0,61 bij een steekproefgrootte van 20 is:

$$t' = \frac{0,61 \cdot \sqrt{20-2}}{\sqrt{1 - 0,3721}} = 3,26$$

De kans op een correlatie van minstens 0,61 bij een steekproefgrootte van 20 is:

$$P \{ t' \geq |3,26|, df = 18 \} = 0,0044$$

Het gemiddeld aantal te verwachten correlaties van minstens 0,61 bij 50 variabelen is bij toepassing van de clusteranalyse:

$$M = \frac{1}{2} \cdot 50 \cdot (50-1) \cdot 0,0044 = 5,39$$

De resultaten hebben dus geen betekenis.

Bij zulke kleine N is het wel zaak ook de hoogte van de correlaties in de beschouwing te betrekken. Bij steekproeven beneden de dertig kan de point-biserial r maar een beperkt aantal waarden aannemen. Hierdoor is het mogelijk bij het gebruik van een overeenkomstindex, die gelijk is aan één van die waarden, een onjuist beeld bij de toetsing te verkrijgen. Zo gaf analyse van simulatie I met een overeenkomstindex van 0,60 13 'significante' correlaties. Hiervan bleken er 7 gelijk aan de overeenkomstindex. Toevallig kunnen er bij een dergelijke grens maar gemiddeld 6 correlaties voorkomen met een standaardafwijking van 2,5. Het gevonden aantal is dus duidelijk meer. Bij uitsluiting van de 7 'significante', die gelijk zijn aan de overeenkomstindex, is het aantal echter precies gelijk aan het gemiddeld aantal te verwachten op grond van toeval.

Simulatie II is geanalyseerd met een veel lagere overeenkomstindex, namelijk 0,25. Dit is mogelijk omdat in dit onderzoek de steekproef veel groter is: $N=100$. De analyse gaf 10 niet-specifieke clusters, waarvan 9 bestaande uit 2 operatoren en 1 uit 3. Totaal waren er dus 11 'significante' correlaties. Van de 10 clusters bleken er 4 een K.R.(21) te hebben die lager was dan de 5% grens: 0,44, 5 een K.R.(21) beneden de 1% grens: 0,56 en 1 een K.R.(21) boven de 1% grens. Dit laatste was het cluster met de drie operatoren. Controle met behulp van de K.R.(21) formule zou dus al sterk het vermoeden hebben bevestigd dat de resultaten zinloos zijn. De toetsing op toeval geeft de volgende uitkomsten (gebruikt wordt de normale benadering van de statistische betrouwbaarheid). De standaardafwijking van de Fisher z-waarde bij een steekproefgrootte van 100 is:

$$\sigma_z = \frac{1}{\sqrt{100-3}} = 0,0101$$

De standaard normale t van een correlatie van 0,25 (z-waarde 0,26) bij een steekproefgrootte van 100:

$$t = \frac{0,26}{0,0101} = 0,26$$

De kans op een correlatie van minstens 0,25 bij een steekproefgrootte van 100:

$$P \{ t \geq |0,26| \} = 0,0094$$

Het gemiddeld aantal te verwachten correlaties van minstens 0,25 bij 50 variabelen is bij toepassing van de clusteranalyse:

$$M = \frac{1}{2} \cdot 50 \cdot (50-1) \cdot 0,0094 = 11,51$$

Elf 'significante' correlaties is dus een duidelijk toevalsresultaat.

6.3. Kritische punten iteratieve clusteranalyse

De iteratieve clusteranalyse is een verbetering van de door Wherry en Gaylord (1943) ontwikkelde iteratieve faktoranalyse. Het belangrijkste verschil tussen deze twee analyses is gelegen in de wijze waarop de clusters worden gevormd. Bij de clusteranalyse gebeurt dit additief. Uitgaande van één operator wordt nagegaan of het cluster met één operator kan worden uitgebreid, met de som van de twee operatoren wordt dit proces herhaald enz. totdat geen één operator voldoende overeenkomstig is met het gevonden cluster. De iteratieve faktoranalyse echter gaat uit van de totale som van alle variabelen, berekent de correlaties van alle variabelen met deze som, vormt een nieuwe som van alle variabelen, die in voldoende mate correleren, berekent weer de correlaties van alle variabelen, met deze tweede totaalscore vormt weer een nieuwe som enz., totdat er geen veranderingen meer optreden.

De analyse van Wherry en Gaylord is duidelijk iteratief van aard, de iteratieve clusteranalyse niet. Een betere naam voor de laatste analyse zou additieve faktoranalyse zijn geweest of naar de methode van Lenz en Whitmer (1941) 'item synonymization'. Ook deze onderzoekers beginnen met één item en bouwen dan een 'synonymy' op. Om echter zowel zijn afkomst, als zijn verschil met faktoranalyse tot uitdrukking te brengen is besloten tot genoemde naam.

De analysemethode van Wherry en Gaylord heeft een aantal bezwaren. De belangrijkste is wel de 'kunstmatigheid' van de correlaties. De variabelen bepalen zelf de som en correleren dus gedeeltelijk met zichzelf. De correlaties, waar de beslissingen opgenomen moeten worden, zijn dus niet alleen afhankelijk van de hoogte van de ladingen, maar ook van het aantal variabelen die de som bepalen zowel als hun ladingen (zie ook paragraaf 4.5.).

Mosier (1936) heeft er al op gewezen, dat het gebruik van 'item-totaal' correlaties moeilijkheden kan geven als

1. de factoren even sterk zijn, en
2. de items geen 'simple-structure' hebben, dat wil zeggen ladingen hebben op meer factoren.

Wherry en Gaylord (1943) vermelden dat het eerste punt hun in een testcase met de iteratieve factoranalyse is overkomen. "In two attempts to use this method empirically the writers found one test which worked out quite smoothly due to unequally represented factors, whereas in a second test the two main factors were equipotent and came out as a single sub-test until S_1 was broken up into S_{1A} and S_{1B} by visual inspection, after which the method produced to work satisfactorily" (p. 263).

Juist door zijn additieve karakter kan bovenstaande niet bij de iteratieve clusteranalyse voorkomen. Hiervoor kan onder andere gewezen worden op simulatie V (zie paragraaf 6.2.1.). De faktorstructuur van dit simulatieonderzoek bestaat uit twee volledig overeenkomstige, elkaar gedeeltelijk overlappende factoren. Bijnen (1968), die de kritiek van Mossier's eerste punt herhaalt, acht het een bezwaar, dat één van de randvariabelen of tussenvariabelen als eerste wordt gekozen. De volgorde van keuze van de operatoren zegt echter weinig. Het gaat om de ladingen. Het komt wel eens voor, dat de eerste variabele uiteindelijk een lading krijgt beneden de overeenkomstindex!

Guilford (1954), die de analyse van Wherry en Gaylord en die van Lentz en Whitmer bespreekt, komt met de kritiek dat "Neither of these methods will necessarily produce univocal tests in the sense that each measures only one common factor" (p. 434). Deze mening samen met het tweede punt van Mossier zijn niet geheel onjuist. Bij de

bespreking van de nauwkeurigheid van de ladingen en de subclusters gaan we hier verder op in.

Tellegen (1968) heeft bezwaar tegen het toetsend gebruik van de clusteranalyse, zoals deze door Hermans (1967, p.39) is uitgevoerd. "We kunnen niet met Hermans meegaan, als hij stelt dat deze methode voor de één-faktorhypothese toetsend zou zijn. Wanneer er namelijk geen statistische grens gesteld wordt bij welk percentage uitval van items de één-faktorhypothese verworpen moet worden, kunnen we niet van toetsing spreken. Van een werkelijke toetsing van deze hypothese is pas sprake wanneer aangetoond is dat de rang van de intercorrelatiematrix één is". Deze kritiek is niet onjuist, al moet er wel bij worden gezegd, dat Hermans nooit de één-faktorhypothese heeft willen toetsen. Getoetst wordt of er sprake is van een algemene faktor. Het is namelijk niet nodig voor een schaal, dat de intercorrelatiematrix van de items de rang één heeft. De opvattingen van Tellegen hierover zijn ons inziens te streng. In paragraaf 3.4. is al uiteengezet, dat een faktorstructuur van één algemene faktor met gebalanceerde groepsfactoren ook adequaat is. Cronbach (1951) heeft verder er al op gewezen, dat kleine afwijkingen in de 'balancing' geen bezwaar zijn, indien er maar een voldoende aantal operatoren zijn. Wij willen wijzen op de resultaten van simulatie IV. In dit onderzoek (zie ook paragraaf 3.6. bijlage II.1. en III) is uitgegaan van een faktorstructuur bestaande uit één algemene faktor en vijf subfactoren bij 27 operatoren. Berekend zijn nu de correlaties van de clusterscores van de 600 elementen van het eerste algemene cluster met de ware scores van alle zes factoren. Deze waarden bleken respectievelijk: 0,929, 0,083, 0,082, 0,002, 0,121 en 0,015 te zijn. De nul-hypothese kan voor de laatste vijf correlaties niet worden verworpen ($P \leq 0,14$).

Hiermee wordt bevestigd, dat kleine niet-gebalanceerde subfactoren een verwaarloosbare invloed hebben op de totaalscore indien er maar voldoende operatoren zijn.

Segers (1965, p.88 en 1968, p.123) heeft een aantal bezwaren tegen de iteratieve clusteranalyse naar voren gebracht, die hierop neerkomen dat de voorwaarden, waaraan de analyse moet voldoen tot een aantal beperkingen en vergrovingen voeren. Dit is in vergelijking met andere technieken wat overdreven. De beperkingen en vergrovingen zijn veelal minder ernstig dan die van andere analyses. Welke beperkingen heeft de clusteranalyse? Ten eerste moeten de varianties van de variabelen zoveel mogelijk gelijk zijn, ten tweede kan het aantal operatoren niet veel groter zijn dan 200 tot 500 (afhankelijk van het computerprogramma) en het aantal elementen niet meer dan 200 tot 1000. Hier staat tegenover, dat bij de hoofdassenanalyse het aantal variabelen niet groter dan 70 tot 150 kan zijn, het aantal proefpersonen altijd gelijk of groter moet zijn dan het aantal variabelen en de maten verdeelt op interval niveau dienen te zijn. De hoofdassenanalyse is dus in aantal opzichten beperkter dan de clusteranalyse, daarbij wordt nog afgezien van een toetsings- of een hiërarchisch-onderzoek, dat met de clusteranalyse vrij eenvoudig is uit te voeren. Welke vergrovingen voert de analyse in? De meest opvallende en ook de belangrijkste is de codering in nul en één. Deze 'vergroving' maakt het mogelijk geen aannamen nodig te hebben ten aanzien van de maten van de variabelen. Hierdoor wordt de gebruiksmogelijkheid van de analyse sterk uitgebreid. Wij wijzen o.a. op de problemen, die zich voordoen, als de onderzoeker een factoranalyse wil uitvoeren op tegelijkertijd kwalitatieve data als beroep en sexe en kwantitatieve data als testuitslagen en leeftijd. Bovengenoemde 'vergroving' heeft wel in sommige gevallen tot gevolg dat de ladingen minder nauwkeurig worden geschat. In de volgende

paragraaf wordt daar verder op ingegaan.

De voorlopige conclusie is dat de gegeven kritieken op zichzelf vaak niet onjuist zijn, echter meestal vrij specifiek zijn geformuleerd, als zou de kritiek alleen voor de clusteranalyse gelden en niet voor bijvoorbeeld de hoofdassenanalyse. Dit zal nader worden onderzocht voor de schattingen van de ladingen en voor de bepaling van de subclusters.

6.3.1. De ladingen

De schatting van de ladingen is voor de iteratieve clusteranalyse een kritisch punt. Er zijn namelijk verschillende redenen, waarom de iteratieve clusteranalyse de ladingen minder nauwkeurig kan geven dan de hoofdassenanalyse. Ten eerste werkt de clusteranalyse met nul en één coderingen en de hoofdassenanalyse met maten in een groot aantal eenheden. Bij normale verdelingen kan dus het z.g. 'coarse grouping' effect (Peters and Van Voorhis, 1941) optreden, waardoor de correlaties van de clusteranalyse te laag worden. Door omzetting van de point-biserial r in de biserial r (zie bijlage I.5.) wordt een correctie aangebracht voor dit effect.

Ten tweede kunnen bij kleine clusters 'part-whole' effecten optreden, waardoor de correlaties te hoog worden. Guilford (1954) en Henrysson (1963) hebben formules ontwikkeld om deze verhoging uit te schakelen bij 'item-total' correlaties. Deze formules zijn bruikbaar bij de clusteranalyse. De formule van Henrysson voor de point-biserial r is

$$r' = (\sqrt{n/n - 1}) \cdot (r_p \cdot \sigma_t - \sqrt{pq}/\sqrt{\sigma_t^2 - \Sigma pq})$$

Hierbij is r' de gecorrigeerde correlatie,

n het aantal operatoren van de cluster,

r_p de berekende point-biserial correlatie,

σ_t de variantie van de clusterscores, en

p de proportie één coderingen van de operator ($q = 1 - p$).

Ten derde kunnen de ladingen van variabelen, die in meerdere clusters voorkomen, worden verlaagd of verhoogd. Dit is vooral het geval bij de variabelen van een subcluster, als deze variabelen in een zelfde algemener cluster voorkomen. Mosier (1936) en Guilford (1954) hebben hier, weliswaar in andere termen, op gewezen.

Het is nu belangrijk om na te gaan in hoeverre bovengenoemde beperkingen ook praktische consequenties hebben. Daarvoor is bij simulatie III t/m VIII het gemiddelde absolute verschil in z-waarden bepaald tussen de gevonden ladingen en de populatie ladingen voor de hoofdassenanalyse en clusteranalyse. Hierbij zijn alleen de factoren bekeken, niet de subfactoren. Bij de hoofdassenanalyse is verder het werkelijk aantal factoren geroteerd. In tabel 6.3.1.1. staan de resultaten van deze vergelijking.

Tabel 6.3.1.1.
Gemiddeld absolute verschil in z-waarden van
de ladingen geschat door de hoofdassenanalyse
en de clusteranalyse bij simulatie III t/m VIII

| | N | n | F | hoofdassen | cluster |
|------|-----|----|---|------------|-------------|
| III | 100 | 30 | 1 | 0,07 | 0,12 (0,10) |
| IV | 600 | 27 | 1 | 0,05 | 0,15 (0,09) |
| V | 100 | 12 | 2 | 0,13 | 0,11 (0,12) |
| VI | 100 | 14 | 3 | 0,14 | 0,16 (0,16) |
| VII | 100 | 14 | 3 | 0,19 | 0,14 (0,17) |
| VIII | 100 | 14 | 3 | 0,15 | 0,11 (0,11) |

Uit tabel 6.3.1.1. is af te lezen, dat als de factoren overgepresenteerd worden en de steekproef groter wordt de hoofdassenanalyse duidelijk nauwkeuriger is dan de clusteranalyse. Bij simulatie III en IV geeft na correctie van de ladingen voor 'coarse grouping' de clusteranalyse resp. 0,10 en 0,09 als gemiddeld absoluut verschil. Bij de andere vier onderzoeken geeft deze correctie geen verbetering, ook niet als

eerst voor het verhogend effect van de 'part-whole' correlaties is gecorrigeerd (zie de tussen haken geplaatste waarden van tabel 6.3.1.1.).

Hierbij dient wel opgemerkt te worden dat de clusteranalyse vaak twee keer is uitgevoerd of omdat uit de resultaten bleek dat een verhoging of een verlaging van de overeenkomstindex een beter resultaat zou geven of om door opgave van een pivot tussenvariabelen beter te presenteren. Zo geeft simulatie V bij een overeenkomstindex van 0,30 ook twee clusters echter met meer specifieke clusters namelijk drie. Hierdoor is de schatting van de ladingen ook minder nauwkeurig namelijk met een verschil van 0,13. Bij een overeenkomstindex van 0,20 is er maar één specifiek cluster en is het gem. absolute verschil 0,11 (zie paragraaf 6.2.1.). Bij simulatie VI moest de derde cluster als pivot worden opgegeven om een betere representatie te geven. Dit gebeurde zonder 'voorkennis' van de populatiestructuur dat wil zeggen alle variabelen werden opgegeven, die in het cluster zaten of een lading boven de overeenkomstindex hadden. Deze bewerking vond ook plaats bij simulatie VII. Zonder deze bewerkingen zijn de gemiddelde verschillen respectievelijk 0,19 en 0,20.

Simulatieonderzoek IV kan nadere informaties geven over subcluster-effecten op de schattingen van de ladingen. De factorstructuur van dit onderzoek bestaat uit één algemeen cluster en vijf subclusters (zie bijlage III). De ladingen van de variabelen in het eerste cluster lopen per drie variabelen op van 0,40 t/m 0,80, dus variabele 1, 2 en 3 hebben een lading van 0,40, variabele 4, 5 en 6 van 0,45, variabele 7, 8 en 9 van 0,50 enz. Vier subclusters bestaan uit vier variabelen met

ladingen van 0,40 t/m 0,60. Een subcluster heeft drie variabelen. Vier subclusters hebben een variabele met een negatieve lading. Iedere variabele komt niet meer dan maximaal in één subcluster voor.

In tabel 6.3.1.2. staat de vergelijking voor de 8 operatoren, die geen lading hebben in de subclusters, de 15 operatoren met een positieve lading en de 4 met een negatieve. Berekend zijn gemiddelde (niet absolute) verschillen tussen de geschatte ladingen in z-waarden en de populatie z-waarden. Bij de clusteranalyse is dit gebeurd met de voor 'coarse grouping' gecorrigeerde ladingen.

Tabel 6.3.1.3.
Gemiddelde (niet absolute) verschillen
in z-waarden van variabelen van simulatie IV met geen, positieve en negatieve ladingen in subclusters bij de hoofdassen- en clusteranalyse.

| | n | hoofdassen | cluster |
|------------------|----|------------|---------|
| geen lading | 8 | + 0,05 | + 0,08 |
| positieve lading | 15 | + 0,06 | + 0,10 |
| negatieve lading | 4 | - 0,03 | + 0,03 |

Bij beide methoden doen de verwachte effecten zich voor. Bij de negatieve ladingen in sterkere mate dan bij de positieve ladingen. De kritiek van Mossier en Guilford valt niet te ontkennen. Het is echter de vraag of hij wel zo specifiek is voor de clusteranalyse.

Samenvattend kan worden gesteld, dat de hoofdassenanalyse bij overpresentatie van de factor ($n \geq 20$) en/of bij grote steekproeven ($N > 500$) nauwkeuriger schattingen geeft dan de clusteranalyse van de ladingen. In de andere gevallen is de clusteranalyse echter niet onnauwkeuriger dan de hoofassenmethode.

6.3.2. Subclusters

De iteratieve clusteranalyse is niet in staat subclusters te bepalen, de hoofdassenanalyse wel. Dit voordeel van de hoofdassenanalyse is echter vrij theoretisch van aard. De factorstructuur van factoren met subfactoren is meestal geen 'simple-structure'. Bij varimax rotatie worden dus de subfactoren versterkt ten koste van de factoren. Tenopyr en Michael (1963, 1964) hebben hier al op gewezen. De onderzoeker moet dus weten uit hoeveel factoren en hoeveel subfactoren de populatiestructuur bestaat en dit aantal moet hij exact weten, anders krijgt hij een resultaat, dat zowel voor de factoren als de subfactoren gefalsificeerd is. Bovengenoemd effect treedt vooral op, als er sprake is van een algemene factor (Tenopyr and Michael, 1963, 1964). Met simulatie VIII zullen we demonstreren dat dit echter ook het geval is bij groepsfactoren. In tabel 6.3.2.1. staat de factorstructuur van simulatie VIII.

Tabel 6.3.2.1.
Factorstructuur simulatie VIII
(nulwaarden zijn weggelaten)

| | F_1 | F_2 | F_3 | F'_1 | F'_2 | F'_3 |
|----|-------|-------|-------|--------|--------|--------|
| 1 | 0,60 | | | | 0,54 | |
| 2 | 0,60 | | | 0,50 | | |
| 3 | 0,40 | | -0,50 | | | |
| 4 | 0,50 | | | 0,45 | | |
| 5 | 0,60 | | | | -0,44 | |
| 6 | 0,80 | | | | | |
| 7 | | 0,60 | 0,44 | | | |
| 8 | | 0,60 | | -0,50 | | |
| 9 | 0,70 | 0,40 | | | | |
| 10 | | 0,50 | | 0,35 | | 0,70 |
| 11 | | 0,60 | | | | |
| 12 | | 0,80 | | | | 0,38 |
| 13 | | | 0,80 | | 0,39 | |
| 14 | | | 0,70 | | | |

Deze structuur bestaat uit drie groepsfactoren en drie subfactoren. In tabel 6.3.2.2. staan de resultaten van de hoofdassenanalyse voor de eerste

factor en wel na extractie; na rotatie van factor 1 en 2; rotatie van factor 1, 2 en 3; 1, 2, 3 en 4; 1, 2, 3, 4 en 5; en 1, 2, 3, 4, 5 en 6. De resultaten zijn gegeven in 100 x de z-waarden. Tevens staat in deze tabel het resultaat van de iteratieve clusteranalyse voor deze eerste factor. Het volledige resultaat van deze analyse voor de eerste drie factoren is gepubliceerd in Boon van Ostade (1965 a en b). Deze resultaten zijn verkregen met enigszins afwijkende keuze van de eerste variabele als in deze publicatie is beschreven.

Tabel 6.3.2.2.

Resultaten in 100x de z-waarden van
de eerste factor, simulatie VIII

| populatie | extractie | | | rotatie | | | iteratieve | |
|-----------|-----------|-----|-----|---------|-----|-----|------------|-----|
| | | | 2 | 3 | 4 | 5 | 6 | |
| 1 | 69 | 55 | 55 | 79 | 69 | -2 | -3 | 65 |
| 2 | 69 | 91 | 95 | 89 | 48 | 41 | 33 | 76 |
| 3 | 42 | 74 | 79 | 44 | 47 | 42 | 19 | 38 |
| 4 | 55 | 79 | 83 | 87 | 51 | 45 | 26 | 65 |
| 5 | 69 | 73 | 68 | 58 | 68 | 126 | 147 | 62 |
| 6 | 110 | 100 | 97 | 116 | 129 | 78 | 55 | 85 |
| 7 | 0 | -5 | -14 | 13 | 20 | 9 | 4 | -7 |
| 8 | 0 | -4 | -12 | -17 | 19 | 25 | 22 | -23 |
| 9 | 87 | 79 | 71 | 89 | 107 | 59 | 59 | 78 |
| 10 | 0 | 18 | 9 | 2 | -9 | -3 | 5 | 5 |
| 11 | 0 | 23 | 13 | 10 | 16 | 18 | 8 | 19 |
| 12 | 0 | 9 | -2 | -17 | 8 | 7 | 13 | -2 |
| 13 | 0 | -32 | -35 | 6 | 5 | -19 | -24 | -11 |
| 14 | 0 | -40 | -42 | -7 | -1 | 11 | 17 | -4 |

Gem absolute 0,17 0,19 0,11 0,11 0,23 0,30 0,10
z-waardeverschil

Aan de gemiddelde absolute z-waardeverschillen is af te lezen, dat varimax rotatie zinvol is (van 0,17 naar 0,11), echter duidelijk ongunstige resultaten geeft als subfactoren worden meegeroteerd (van 0,11 naar 0,30). Niet veel aandacht moet worden besteed aan het gelijke resultaat van de eerste factor bij rotatie van de drie groepsfactoren (3 : 0,11) en de drie groepsfactoren en de eerste subfactor (4 : 0,11). Deze rotatie gaat in dit geval niet ten koste van de eerste groepsfactor, echter wel van de tweede en derde.

Hetzelfde effect treedt ook op bij de subfactoren. In tabel 6.3.2.3. is dit nader gedemonstreerd voor de eerste subfactor.

Tabel 6.3.2.3.

Resultaten in 100x de z-waarden van
de eerste subfactor, simulatie VIII

| populatie | extractie | | rotatie | | |
|----------------------------------|-----------|-----|---------|------|------|
| | | | 4 | 5 | 6 |
| 1 | 0 | -4 | 30 | 17 | 16 |
| 2 | 55 | 42 | 87 | 89 | 93 |
| 3 | 0 | -25 | 4 | 3 | 9 |
| 4 | 48 | 31 | 74 | 78 | 87 |
| 5 | 0 | -21 | -5 | 5 | 8 |
| 6 | 0 | -27 | 12 | 12 | 17 |
| 7 | 0 | 6 | -16 | -16 | -13 |
| 8 | -55 | -65 | -113 | -107 | -102 |
| 9 | 0 | -24 | -3 | -5 | -3 |
| 10 | 37 | 60 | 15 | 16 | 14 |
| 11 | 0 | 12 | -18 | -17 | -15 |
| 12 | 0 | 14 | -29 | -28 | -29 |
| 13 | 0 | 3 | 2 | -1 | -1 |
| 14 | 0 | -9 | -12 | -6 | -6 |
| Gem absolute z-waardeverschil | 0,14 | | 0,20 | 0,18 | 0,18 |

Subclusters of subfactoren zijn alleen te ontdekken als het aantal factoren exact bekend is. Dit is mogelijk door b.v. eerst een clusteranalyse uit te voeren en daarna een hoofdassenanalyse. Een dergelijke methode is echter alleen betrouwbaar bij grote steekproeven ($N > 500$) en/of bij voldoende presentatie van de factoren door variabelen ($n \geq 20$). In alle andere gevallen kunnen de data beter gecorrigeerd worden voor de invloed van de (algemene) factor (Boon van Ostade 1965, b), zoals o.a. is geschiedt door Reumkens (1962) Gooyen (1966) en Jansen van der Sligte (1969). De clusteranalyse wordt dan herhaald op z.g. ipsatieve scores, gevormd door aftrekking van de clusterscore van de (algemene) factor van alle scores van één proefpersoon of element bij die operatoren die een lading hebben in de (algemene) factor.

7. Samenvatting

Op logische gronden kan worden gesteld, dat een klassificatie, als klassificatie van een verzameling elementen, doelmatiger is

- a. naarmate er meer kenmerken opgaan voor de elementen die in de klasse vallen,
- b. naarmate deze kenmerken niet opgaan voor de elementen die niet in de klasse vallen, en
- c. naarmate vijftig procent van de elementen in de klasse valt.

Een klassificatie is in een empirische wetenschap verder als klassificatie ondoelmatig wanneer

- a. er geen operator is om de klassificatie uit te voeren,
- b. alle elementen van de verzameling in de klasse vallen, of
- c. geen één element van de verzameling in de klasse valt.

Een operator is hierbij gedefinieerd als een variabele, die met behulp van een vast observatie- en coderingsvoorschrift onderscheidingen aanbrengt tussen de elementen. Onder elementen worden daarbij verstaan: individuen of stimuli.

De waarde van een operator voor een bepaalde klassificatie wordt primair gegeven door de validiteit van de operator, dat wil zeggen door de kans, dat bij toepassing van de operator constante- en/of toevallige fouten worden gemaakt. In de analyse wordt deze validiteit bepaald door berekening van de correlatie coëfficiënt tussen de operator en de bedoelde klassificatie.

De bedoelde klassificatie kan worden geconstrueerd door explicatie van de eigenschap in een aantal nevenschikkende kenmerken, die zodanig in operatoren worden gespecificeerd dat de operatoren variëren in hun niet-relevante aspecten. Hierdoor is de representativiteit van de set van operatoren voor de bedoelde klassificatie

theoretisch gewaarborgd. Met behulp van de verhouding tussen het aantal operatoren dat het element tot de klasse rekent ten opzichte van het totaal aantal operatoren, waarmee het element is geobserveerd voor de bedoelde klassificatie, kan de klassificatiemaat van het element worden bepaald.

De analyse dient nu om de set van operatoren empirisch op bovenstaande veronderstellingen te controleren. Dit is mogelijk doordat gesteld kan worden, dat de centroïd factorstructuur isomorf is met bovenstaande theorie. Een centroïde factorstructuur bestaat uit één algemene factor (de bedoelde conditie), een aantal gebalanceerde groepsfactoren (de niet-relevante condities) en uit n (unieke) factoren, waarbij n het aantal operatoren is. Een set van operatoren is nu te toetsen door na te gaan in hoeverre de intercorrelaties van deze operatoren een dergelijke factorstructuur vormen. Voor een dergelijke toetsing is het nodig van tevoren de minimum validiteit of overeenkomstindex te bepalen en wel om de convergerende validiteit van de set te kunnen controleren. De divergerende validiteit kan alleen worden gecontroleerd als in het onderzoek variabelen van niet-bedoelde condities zijn opgenomen.

De analyse wordt meestal niet toetsend toegepast, maar explorerend. In het laatste geval tracht de onderzoeker clusters te vinden, die voldoen aan de eisen van de convergerende- en divergerende validiteit. Bij controlerend-toetsend- en controlerend-exploratief onderzoek vindt de analyse plaats op een van tevoren bepaalde set of pivot van operatoren, waarbij in het eerste geval getracht wordt het beste cluster te vinden, terwijl in het tweede geval de intercorrelaties van de pivot worden geanalyseerd conform de exploratieve methode.

De waarde van de geconstrueerde klassificaties of clusters kan worden bepaald met behulp van de formules 20 en 21 van Kuder en

Richardson. Daarbij dient echter wel rekening te worden gehouden met het feit, dat deze formules te hoge waarden geven.

Uitvoering van de analyse bij minder dan 60 elementen blijkt de resultaten duidelijk ongunstig te beïnvloeden. Verder blijken specifieke clusters, bestaande uit één operator, van weinig betekenis. De operatoren dienen met nul en één te worden gecodeerd met zoveel mogelijk gelijke p-waarden (kans op coderingen één) van 50%. Wanneer bij de data-verzameling de observaties in zoveel mogelijk (ordinale) eenheden zijn gecodeerd, roept dit voorschrift geen ernstige bezwaren op. Voor eenvoudige factorstructuren, bestaande uit weinig, niet-overlappende factoren, kan van de regel van gelijke p-waarden worden afgeweken. De van tevoren gestelde overeenkomstindex bepaalt ook de orde, waarin de klassificatie plaats vindt. Door gebruik van verschillende overeenkomstindices is een onderzoek naar de hiërarchie van de klassificatie mogelijk.

De vorming van clusters vindt in de analyse plaats door 'item synomization'. Van één operator wordt uitgegaan en dit cluster wordt uitgebreid met steeds één operator, afhankelijk van de correlaties met de ongewogen somscore van het cluster. De operator, die wordt opgenomen, heeft de hoogste correlatie boven of gelijk aan de gestelde overeenkomstindex. De gevormde clusters dienen te worden gecontroleerd op

- a. hun kans op ontstaan,
- b. hun betrouwbaarheid, en
- c. hun homogeniteit van clusterladingen.

De analyse kan worden toegepast als R- en Q-analyse bij individuen en stimuli. Het verschil tussen R- en Q-analyse is gelegen in de berekening van de correlaties. In het eerste geval gebeurt dit

tussen de operatoren over de elementen, in het tweede geval tussen de elementen over de operatoren. Bij Q-analyse zijn dus de 'elementen' de operatoren. Hierdoor wordt het mogelijk onderzoek te doen van betekenissen (Q-analyse van stimuli), van psychische functies (R-analyse van stimuli), van interindividuele verschillen (R-analyse van individuen) en van typeringen (Q-analyse van individuen). Verder is de analyse bruikbaar voor conditie-, hiërarchisch- en/of predictief-onderzoek.

De vergelijking met de hoofdassenanalyse (uitgevoerd met communaliteiten van één en varimax rotatie van factoren met eigenwaarden groter dan één) heeft plaats gevonden met acht gesimuleerde datamatrices voor verschillende factorstructuren. Uit de vergelijking blijkt, dat de iteratieve clusteranalyse betrouwbaarder is ten aanzien van de bepaling van het aantal factoren. De hoofdassenanalyse daarentegen geeft nauwkeuriger schattingen van de ladingen bij grote steekproeven van elementen (vijfhonderd of meer) en/of bij overpresentatie van de factoren (met meer dan twintig variabelen).

Geconcludeerd kan worden, dat de iteratieve clusteranalyse als klassificatie methode goed bruikbaar is in die gevallen waar

- a. veel variabelen (of, voor Q-analyse, veel elementen) moet worden geanalyseerd,
- b. de steekproef van elementen (of, voor Q-analyse, van variabelen) betrekkelijk klein is,
- c. kwalitatieve en/of ordinale data zijn verzameld, en/of
- d. conditie- en/of type-onderzoek wordt gevraagd.

8. Summary

It may be stated on logical grounds that the efficiency of a classification as a classification of a set of elements increases

- a. as the number of characteristics that may be applied to assign the elements to the class increases,
- b. as these characteristics are not applicable to the elements that are excluded from the class, and
- c. the proportion of the elements which fall into the class approached 50%.

Furthermore a classification as a classification in an empirical science is inefficient when

- a. no operator exists to effect the classification,
- b. all the elements in the set fall into the class, or
- c. not one of the elements in the set fall into the class.

An operator is defined here as a variable which with fixed observation and coding rules establishes a differentiation among the elements. Elements are defined here as either individuals or stimuli.

The value of an operator for a specific classification is given primarily to the validity of the operator i.e. in the probability that in the application of the operator errors (constant and/or variable) will be made. This validity is determined analytically by calculating the correlation coefficient between the operator and the classification referred to.

This classification can be constructed through the explication of the quality by the use of a number of co-ordinating characteristics, which are specified in the operators in such a way that the operators vary in their non-relevant aspects. Through this the representativeness of the set of operators for the classification that we are concerned with, is

theoretically established. The relation between the number of operators which assign the element to the class and the total number of operators used to observe the element for the classification referred to, gives the measure of classification for the element.

The analysis now serves to examine a set of operators empirically on the above assumptions. We can do this because it can be stated that a centroid factor structure is isomorphic with the above theory. A centroid factor structure consists of one general factor (the relevant condition), a number of balanced group factors (the non-relevant conditions) and of \underline{n} unique factors where \underline{n} is the number of operators. A set of operators can now be tested by determining the extent to which the inter-correlations of these operators form a similar factor structure. For such testing it is necessary to determine beforehand the minimum validity or criterium of similarity in order to test the convergent validity of the set. The divergent validity can only be tested if variables of non-relevant conditions are included.

The analysis is not usually applied to verify but to explore. As regards the latter the researcher tries to find clusters that satisfy the conditions of the converging and diverging validity. The analysis may be carried out through controlled-testing and controlled-explorative investigations on a previously specified set or pivot of operators where in the first case the best cluster is sought, whereas in the second case the inter-correlations of the pivot are analysed in conformity with the exploratory method.

The worth of the constructed classifications or clusters can be specified using Kuder and Richardson's formulas 20 and 21. Attention should be called to the fact that values given by these formulas are too high.

If the analysis is carried out with fewer than 60 elements the results are clearly effected unfavourably. Furthermore specific clusters which consists of only one operator turn out to have little significance. The operators should be coded with zero and one, with as many possible equal p-values (probability of coding one) of 50%. This coding rule raises no serious objections in data compilation when observations are coded in as many (ordinal) units as possible. The rule of equal p-values does not need to be followed in the case of simple factor structures consisting of few non-overlapping factors. The pre-determined criterium of similarity also determines the order in which classification takes place. An analysis can be made of the hierarchy of the classification by using different criteria of similarity.

Cluster formation takes place in the analysis through "item synomization". One operator is taken as a starting-point and this cluster is increased by one operator at a time, dependent on the correlation with the unweighted sum score of the cluster. The operator to be selected is the one with the highest correlation coefficient above or equal to the specified criterium of similarity. The clusters formed should be tested for

- a. the probability that they will exist,
- b. their reliability, and
- c. the homogenity of their cluster-loadings.

The analysis can be carried out through R- and Q-analysis of individuals and stimuli. The distinction between R- and Q-analysis lies in the computation of the correlation coefficients. The former uses the operators over the elements and the latter the elements over the operators. In Q-analysis therefore, the "elements" may be regarded as the operators. In this way it becomes possible to investigate "meanings" (Q-analysis of stimuli) psychic functioning (R-analysis of stimuli) individual differences

(R-analysis of individuals) and individual typifications (Q-analysis of individuals). In addition the analysis is particularly suitable for conditional, hierarchical and predictive research.

A comparison with the principal axis analysis (performed with communalities of one and varimax rotation of factors with eigenvalues greater than one) has been carried out with eight simulated data matrices of varying factor structures. The comparison shows that iterative cluster analysis is more reliable in specifying the total number of factors. Principal axis analysis on the other hand gives a more precise estimate of the loadings when there is over-presentation of the factors (more than twenty variables).

It may be concluded that iterative cluster analysis is especially useful as a classification procedure when either:

- a. many variables (or, for Q-analysis, many elements) have to be analysed,
- b. the sample of elements (or, for Q-analysis, of variables) is relatively small,
- c. qualitative and/or ordinal data are involved, and/or
- d. conditional and/or type analysis is required.

LITERATUUR

- ADAMS, E., FAGOT, R.F. and ROBINSON, R.E. 1965
A theory of appropriate statistics
Psychometrika, 30, 99-128
- BLALOCK, H.M. 1964
Causal inferences in non-experimental research
Chapel Hill: Un. of North Carolina Press
- BOON VAN OSTADE, A.H. 1965a
Factoranalyse programma's UNIVAC III
In: Personeelsonderzoek met de UNIVAC III
Heerlen: N.V. Ned. Staatsmijnen, Rapport 1628 C.P., 23-35
- BOON VAN OSTADE, A.H. 1965b
De ontwikkeling van enige methodieken
Amsterdam: NIPP-mededelingen no. 165, 8-11
- BOON VAN OSTADE, A.H. 1968
Motivatie- en satisfactie-onderzoek in het bedrijf
Ned. T. Psychol., 23, 347-370
- BOON VAN OSTADE, A.H. en HERMANS, H.J.M. 1967
Een methode ter verhoging van de stabiliteit van
multiple regressieformules
Ned. T. Psychol., 22, 696-715
- BOON VAN OSTADE, A.H. en HERMANS, H.J.M. 1968
Nogmaals over multiple regressieformules
Ned. T. Psychol., 23, 179-183
- BRIDGES, C.C. 1966
Hierarchical cluster analysis
Psychol. Reports, 18, 851-854
- BROVERMAN, D.M. 1961
Effects of score transformations in Q and R factor analysis
techniques
Psychol. Rev., 68, 68-80
- BROWNE, M.W. 1968a
A note on lower bounds for the number of factors
Psychometrika, 33, 233-236
- BROWNE, M.W. 1968b
A comparison of factor analytic techniques
Psychometrika, 33, 267-334

- BRYDEN, M.P. 1960
A non-parametric method of item and test scaling
Educ. Psychol. Measmt., 20, 311-315
- BUROS, O.K. 1961
Tests in print
New Jersey: The Gryphon Press
- BURT, C. 1941
The factors of mind
New York: MacMillan
- BURT, C. 1949
The structure of the mind: A review of the results of
factor analysis
Brit. J. Educ. Psychol., 19, 100-111, 176-199
- BURT, C. 1950
The factorial analysis of qualitative data
Brit. J. Psychol. Stat. Sect., 3, 166-185
- BIJNEN, E.J. 1969
Cluster-analyse. Overzicht en evaluatie van technieken
Tilburg: academisch proefschrift
- CAMPBELL, D.T. and FISKE, D.W., 1959
Convergent and discriminant validation by the Multitrait-
Multimethod matrix
Psychol. Bull., 56, 81-105
- CATTELL, R.B. 1952
Factor analysis
New York: Harper
- CATTELL, R.B. 1966
The scree test for the number of factors
Multiv. Beh. Research, 1, 245-276
- COMREY, A.L. and AL AHUMADA 1964
An improved procedure and program for minimum residual
factor analysis
Psychol. Reports. 15, 91-96
- COOK, S.W. and SELLTIZ, Cl. 1964
A multiple-indicator approach to attitude measurement
Psychol. Bull., 62, 36-55

- COOMBS, C.H. 1964
A theory of data
New York: Wiley
- COAN, R.W. 1964
Facts, factors and artifacts: the quest for psychological meaning
Psychol. Rev., 71, 123-140
- CRONBACH, L.J. 1951
Coefficient alpha and the internal structure of tests
Psychometrika, 16, 297-334
- CRONBACH, L.J. and MEEHL, P.E. 1955
Construct validity in psychological tests
Psychol. Bull., 52, 281-301
- CRONBACH, L.J., RAJARATNAM, N. and GLEESER, G.L. 1963
Theory of generalizability: a liberalization of reliability theory
Brit. J. Statist. Psychol., 16, 1-163
- CURETON, E.E. 1956
Rank-biserial correlation
Psychometrika, 21, 287-290
- DAS GUPTA, S. 1960
Point biserial correlation coefficient and its generalizations
Psychometrika, 25, 393-408
- DAVIS, F.B. 1965
Item analyse
Leuven: Nauwelaerts (vert. Willemaers)
- DAVIS, F.B. and FIFER, G. 1959
The effect on test reliability and validity of scoring
aptitude and achievement tests with weights for every choice
Educ. Psychol. Measm., 19, 159-170
- DIRKZWAGER, A. 1966
Intelligentie en schoolprestaties
Amsterdam: Swets en Zeitlinger
- DUYKER, H.C.J. 1959
Nomenclatuur en systematiek der psychologie
Ned. T. Psychol., 14, 176-217
- DRENTH, P.J. 1968
De psychologische test
Arnhem: Van Loghum Slaterus

- EBEL, R.L. 1962
Content standard test scores
Educ. Psychol. Measmt., 22, 15-25
- ELSHOUT, J.J. 1963
Programma hierarchische clusteranalyse
Amsterdam: Psychol. Lab. G.U.
- EYSENCK, H.J. 1947
Dimensions of personality
London: Routledge and Kegan Paul
- EYSENCK, H.J. 1951
The organization of personality
J. Pers., 20, 101-117
- FELDT, L.S. 1965
The approximate sampling distribution of Kuder-Richardson
reliability coefficient twenty
Psychometrika, 30, 357-370
- FISHER, R. 1932
Statistical methods for research workers
Edinburgh: Oliver and Boyd
- FLANAGAN, J.C. 1962
Discussion
Educ. Psychol. Measmt., 22, 35-39
- FOKKEMA, S.D. 1960
Methodologie en "voorwetenschappelijkheid" in de psychologie
Ned. T. Psychol., 15, 266-294
- FOKKEMA, S.D. 1967
Attributen-correlatie en afhankelijkheidsrelatie in de psychologische
research
Ned. T. Psychol., 22, 568-582
- FREDERIKSEN, N. and MESSICK, S. 1959
Response set as a measure of personality
Educ. Psychol. Measmt., 19, 137-157
- GARDNER, E.F. 1962
Normative standard scores
Educ. Psychol. Measmt., 22, 7-14

- GEER, J.P. VAN DE, 1957
A psychological study of problem solving
Haarlem: De Toorts
- GENGERELLI, J.A. 1963
A method for detecting subgroups in a population and specifying
their membership
J. Psychol., 55, 457-468
- GLASS, G.V. and TAYLOR, P.A. 1966
Factor analytic methodology
Rev. Educ. Research, 36, 566-587
- GOOYEN, J.M.B. 1966
De analyse van een arbeidscurve
Heerlen: afd. Personeelsresearch N.V. Ned. Staatsmijnen,
Rapport no. 60
- GROOT, A.D. DE, 1961
Methodologie
's-Gravenhage: Mouton
- GROOT, A.D. DE, 1966
Vijven en zessen
Groningen: Wolters
- GUILFORD, J.P. 1954
Psychometric methods
New York: McGraw-Hill
- GUILFORD, J.P. 1965
Fundamental statistics in psychology and education
New York: McGraw-Hill, 1950
Intern. Student Edition Tokyo: Kōgakusha
- GULLIKSEN, H. 1950
Theory of mental tests
New York: Wiley
- HAGGARD, E.A. 1958
Intraclass correlation and the analysis of variance
New York: Dryden Press
- HARMAN, H.H. 1960
Modern factor analysis
Chicago: Univ. of Chicago Press

HAZEWINKEL, A. 1968a

De waarde van predictieformules voor de voorspelling
van een criterium

Ned. T. Psychol., 23, 109-121

HAZEWINKEL, A. 1968b

Enige opmerkingen naar aanleiding van "nogmaals
over multiple regressieformules"

Ned. T. Psychol., 23, 184-187

HAVIGHURST, R.J. and MACDONALD, D.V. 1959

Development of the ideal self in New Zealand and American
children

J. Educ. Res., 49, 263-273

HENRYSSON, S. 1950

The significance of factor loadings Lawley's test
examined by artificial samples

Brit. J. Psychol., Stat. Sect., 3, 159-165

HENRYSSON, S. 1963

Correction of item-total correlations in item analysis
Psychometrika, 28, 211-218

HECKE, A.F. VAN, 1965

Psychologisch onderzoek van de houtersfunctie
in afbouw en steenwerk

Heerlen: N.V. Ned. Staatsmijnen, afd. Pers. Research

HERMANS, H. 1967

Motivatie en prestatie

Amsterdam: Swets en Zeitlinger

HERMANS, H. 1968

Prestatie Motivatie Test

Amsterdam: Swets en Zeitlinger

HETTEMA, P.J. 1967

Trekken, processen en persoonlijkheidstests

Ned. T. Psychol., 22, 618-641

HOFSTEE, W.K.B. 1963

Enkele notities over G.J.S. Wilde's "Neurotische labiliteit,
gemeten volgens de vragenlijstmethode"

Ned. T. Psychol., 18, 481-499

HOFSTEE, W.K.B. 1967

Method effects in absolute and comparative judgement
Groningen: van Denderen

HOEL, P.G. 1962

Introduction to mathematical statistics
New York: John Wiley

HORN, J.L. 1965

A rationale and test for the number of factors in
factor analysis
Psychometrika, 30, 179-185

HOTELLING, H. 1933

Analysis of a complex of statistical variables into
principal components
J. Educ. Psychol., 24, 417-441, 498-520

HORN, J.L. 1967

On subjectivity in factor analysis
Educ. Psychol. Measmt., 27, 811-820

HOWARD, K.J. and GORDON, R.A. 1963

Empirical note on the number of factors problem in
factor analysis
Psychol. Reports, 12, 247-250

HUMPHREYS, L.G. 1964

Number of cases and number of factors: an example
where N is very large
Educ. Psychol. Measmt., 24, 457-466

HYMAN, R. 1967

Het psychologisch onderzoek
Utrecht: Spectrum, prisma no. 46 (vert. L.A.M. Timmermans)

JANSEN VAN DER SLIGTE, H.W. 1969

Criteriumanalyse. Fasen in de konstruktie van een empirisch
criterium met behulp van personeelsbeoordelingen
Nijmegen: doctoraalscriptie K.U.

JENKINS, W.L. 1955

An improved method for tetrachoric r
Psychometrika, 20, 253-258

JOHNSON, S.C. 1967

Hierarchical clustering schemes
Psychometrika, 32, 241-254

JONES, L.V. and FISKE, D.W. 1953

Models for testing the significance of combined results
Psychol. Bull., 50, 375-382

JÖRESKOG, K.G. 1963

Statistical estimation in factor analysis

Stockholm: Almqvist and Wiskell

KAHNEMAN, D. and GHISELLI, E.E. 1962

Validity and nonlinear heteroscedastic models

Pers. Psychol., 15, 1-11

KAISER, H.F. 1958

The varimax criterion for analytic rotation in factor analysis

Psychometrika, 23, 187-200

KEATS, J.A. and LORD, F.M. 1962

A theoretical distribution for mental test scores

Psychometrika, 27, 59-72

KEMPEN, G. en BOON VAN OSTADE, A.H. 1969

Een typologie van ideaalbeelden van Europese jeugdigen door
middel van de iteratieve clusteranalyse

Ned. T. Psychol., 24, 46-60

KONINGH, H.L. DE, 1968

Enkele ervaringen met de A.B.V. bij leerling-verplegenden

De Psychol., 3, 306-312

LAWLEY, D.N. 1940

The estimation of factor loadings by the method of maximum
likelihood

Proc. Royal Soc. Edinburgh, Section A, 60, 64-82

LAWLEY, D.N. 1941

Further investigations in factor estimation

Proc. Royal Soc. Edinburgh, Section A, 61, 176-185

LAWLEY, D.N. and SWANSON, Z. 1954

Tests of significance in a factor analysis of artificial data

Brit. J. Stat. Psychol., 7, 75-79

LINDQUIST, E.F. 1953

Design and analysis of experiments in psychology and education

Barton: Houghton Mifflin

LENTZ, T.F. and WHITMER, E.F. 1941

Item synonymization: a method for determining the total meaning
of pencil-paper reaction

Psychometrika, 6, 131-139

- LEYERLY, S.B. 1959
Significance levels for the K.R.(21)
Educ. Psychol. Measmt., 19, 73-75
- LOEVINGER, J. 1947
A systematic approach to the construction and evaluation
of tests of ability
Psychol. Mon., 61, 15-49
- LOEVINGER, J. 1957
Objective tests as instruments of psychological theory
Psychol. Reports, 3, 635-694
- LORD, F.M. 1955
Sampling fluctuations resulting from the sampling of test items
Psychometrika, 20, 1-22
- LORD, F.M. and NOVICK, M.R. 1968
Statistical theories of mental testscores
Reading (Mass.): Addison-Wesley
- LORR, M. and RADHAKRISHNAM, B.K. 1967
A comparison of two methods of cluster analysis
Educ. Psychol. Measmt., 27 I, 47-53
- LORR, M., KLETT, C.J. and McNAIR, D.M. 1963
Syndromes of psychosis
New York: MacMillan
- LUTTE, G., MÖNKS, F., KEMPEN, G., SARTI, S. 1969
Ideaalbeelden van de Europese jeugd. Het onderzoek in Nederland
en Vlaanderen
's-Hertogenbosch: Malmberg
- LUTTE, G. et SARTI, S. 1969
Le moi-idéal de l'adolescent recherche génétique, différentielle
et interculturelle dans sept nations d'Europe
Leuven: Nauwelaerts (publikatie in voorbereiding)
- McNEMAR, Q. 1942
On the number of factors
Psychometrika, 7, 9-18
- McNEMAR, Q. 1963
Psychological statistics
New York: Wiley

McQUITTY, L.L. 1957

A pattern analysis of descriptions of "best" and "poorest"
compared with factor analytic results

Psychol. Mon., 71, nr. 446

McQUITTY, L.L. 1960a

Hierarchical linkage analysis for the isolation of types

Educ. Psychol. Measmt., 20, 55-67

McQUITTY, L.L. 1960b

Hierarchical syndrome analysis

Educ. Psychol. Measmt., 20, 293-304

McQUITTY, L.L. 1960c

Comprehensive hierarchical analysis

Educ. Psychol. Measmt., 20, 805-816

McQUITTY, L.L. 1963

Best classifying every individual at every level

Educ. Psychol. Measmt., 23, 337-346

McQUITTY, L.L. 1964

Capabilities and improvements of linkage analysis as a
clustering method

Educ. Psychol. Measmt., 24, 441-456

McQUITTY, L.L. 1965

A conjunction of rank order typal analysis and item selection

Educ. Psychol. Measmt., 25, 949-962

McQUITTY, L.L. 1966a

Similarity analysis by reciprocal pairs for discrete and
continuous data

Educ. Psychol. Measmt., 26, 825-831

McQUITTY, L.L. 1966b

Single and multiple hierarchical classification by reciprocal
pairs and rank order types

Educ. Psychol. Measmt., 26, 253-266

MILL, J.S. z.j.

A system of logic

London: Longmans, 1949 (heruitgave)

- MOLENAARS, F.M.Th.A. en JANSSEN VAN DER SLIGTE, H.W. 1968
Enquête onder de zusters van Schijndel
Eindhoven: Eindh. Psychol. Inst.
- MOSIER, C. 1936
A note on item analysis and the criterion of internal consistency
Psychometrika, 1, 275-282
- MULDER, M. 1958
Groepsstructuren, motivatie en prestatie
Amsterdam: C.O.P.
- NEDERLANDS INSTITUUT VAN PSYCHOLOGEN, 1969
Documentatie van tests en testresearch in Nederland
Amsterdam: N.I.P.-bureau
- NOVICK, M.R. and LEWIS, Ch. 1967
Coefficient alpha and the reliability of composite measurements
Psychometrika, 32, 1-13
- OERS, J.J.A. VAN, 1964
Een vergelijkend onderzoek van de Amsterdamse Biografische
Vragenlijst in selectie-situatie en niet-selectie-situatie
Heerlen: Staatsmijnen in Limburg, afd. Personeelsresearch
- OPPENHEIM, A.N. 1966
Questionnaire design and attitude measurement
London: Heineman
- PARKINSON, C.N. 1963
De derde wet van Parkinson
Amsterdam: Scheltema en Holkema (vert. J.G.M. Wellen)
- PAYNE, W.H. and ANDERSON, D.E. 1968
Significance levels for the Kuder-Richardson twenty: an automated
sampling experiment approach
Educ. Psychol. Measmt., 28, 23-39
- PEARSON, E.S. and HARTLEY, H.O. 1954
Biometrika tables for statisticians, Volume I
Cambridge: Univ. Press
- PENNINGS, H.J.V. 1966
Beambten en arbeiders in de industrie
Sittard: Alberts
- PENNEL, R. 1968
Iterated communalities, convergence and the number of common factors
Educ. Psychol. Measmt., 28, 793-802

PERRY, N.C. and MICHAEL, W.B. 1954

A tabulation of the fiducial limits for the point biserial correlation coefficient

Educ. Psychol. Measmt., 14, 715-721

PETERS, C.C. and VAN VOORHIS, W.R. 1941

Statistical procedures and their mathematical bases

New York: McGraw-Hill

RICHARDSON, M.W. 1941

The combination of measures

In: P. Horst, The prediction of personal adjustment

New York: Social Science Research Council, Bulletin 48, 379-401

ROSKAM, E.E.Ch.I. en HORSTEN, A. 1968

Algemeen programma factoranalyse

Nijmegen: Psychol. Lab. Programma-Bull. 10

REUMKENS, P.J.M. 1962

Een vergelijkend onderzoek van beoordelingsgegevens, personalia en testgegevens

Geleen: N.V. Ned. Staatsmijnen

RYAN, T.A. 1960

Significance tests for multiple comparison of proportions, variances and other statistics

Psychol. Bull., 57, 318-328

RIJKSARBEIDSBUREAU 1956

Beroepsbeelden

Den Haag: Min. v. Sociale Zaken en Volksgezondheid

RIJKSARBEIDSBUREAU 1958

Beroepeninventarisatie (nieuwe reeks)

Den Haag: Min. v. Sociale Zaken en Volksgezondheid

SAUNDERS, D.M. 1956

Moderator variables in prediction

Educ. Psychol. Measmt., 16, 209-222

SAWREY, W.L., KELLER, L. and CONGER, J.J. 1960

An objective method of grouping profiles by distance functions and its relation to factor analysis

Educ. Psychol. Measmt., 20, 651-673

- SCHEERDER, A. en KERKHOFF-MEEUWIS, I. 1969
Rapport gerichtheidsonderzoek LHNO
Eindhoven: Eindh. Psychol. Inst.
- SEGRS, J.H.G. 1965
Verslag over de interviewphase van het onderzoek naar de
fluctuaties in het arbeidersbestand der nederlandse steenkool-
mijnen
Tilburg: Instituut voor Arbeidsvraagstukken
- SEGRS, J.H.G. 1968
Het personeelsverloop in het ondergronds mijnbedrijf, methoden
en resultaten
Tilburg: academisch proefschrift
- SNEDECOR, C.W. 1937
Statistical methods
Ames (Iowa): Collegiate
- SOKAL, R.R. 1959
A comparison of five tests for completeness of factor extraction
Transactions of the Kansas Academy of Science, 62, 141-152
- STARREN, J.C.M.G.M. 1965
Het meten van studie-motivatie respectievelijk studie-attitude
bij technische vakscholieren
Heerlen: N.V. Ned. Staatsmijnen, afd. Personeels Research
- STEPHENSON, W. 1936
The inverted factor technique
Britt. J. Psychol., 26, 344-361
- STEPHENSON, W. 1953
The study of behavior
Chicago: Univ. of Chicago Press
- STEVENS, S.S. 1951
Handbook of experimental psychology
New York: Wiley
Hoofdstuk 1: Mathematics, measurement and psychophysics
- STEVENS, S.S. 1962
Measurement, psychophysics and utility
In: Churchman, C.W, and Ratoosh, Ph. Measurement: definition and
theories
New York: Wiley, 18-63

- STOUFFER, S.A. GUTTMAN, L., SUCHMAN, E.A., LAZARSFELD, P.F.,
STAR, S.A., CLAUSEN, J.A. 1950
Measurement and prediction
Princeton (New York): Princeton Univ. Press
- TAYLOR, E.K., SCHNEIDER, D.E., and SYMONS, N.A. 1953
A short forced-choice evaluation form for salesmen
Pers. Psychol., 6, 393-401
- TARSKI, A. 1964
Inleiding tot de logica
Amsterdam: Noord-Hollandsche Uitg. Maatsch. (vert. E.W. Beth)
- TELLEGEN, B. 1968
Over rigiditeit
Zaltbommel: Avanti
- TENOPYR, M.L. and MICHAEL, W.B. 1963
A comparison of two computer-based procedures of orthogonal
analytic rotation with a graphical method when a general
factor is present
Educ. Psychol. Measmt., 23, 587-597
- TENOPYR, M.L. and MICHAEL, W.B. 1964
The development of a modification in the normal varimax method
for use with correlation matrices containing a general factor
Educ. Psychol. Measmt., 24, 677-699
- THORNDIKE, R.L. 1953
Who belongs in the family?
Psychometrika, 18, 267-271
- THURSTONE, L.L. and CHAVE, E.J. 1929
The measurement of attitude
Chicago: Univ.Chicago Press
- THURSTONE, L.L. 1947
Multiple factor-analysis
Chicago: Univ.Chicago Press
- TORGERSON, W.S. 1958
Theory and methods of scaling
New York: John Wiley
- TRAVERS, R.M.W. 1955
Educational measurement
New York: MacMillan

- TRYON, R.C. 1957
Reliability and behavior domain validity
Psychol. Bull., 54, 229-249
- TRYON, R.C. 1958
General dimensions of individual differences: cluster analysis
vs multiple factor analysis
Educ. Psychol. Measmt., 18, 477-495
- TUCKER, L.R. 1955
The objective definition of simple structure in linear factor
analysis
Psychometrika, 20, 209-225
- VERNON, P.E. 1950
The structure of human abilities
London: Methuen
- WARD, J.H. 1963
Hierarchical grouping to optimize an objective function
J. Am. Stat. Ass., 58, 236-244
- WASSENAAR, G.M.C. 1956
Labiliteit as temperamentsfactor
Stellenbosch (Z.Afrika): academisch proefschrift
- WELTEN, V.J. 1969
Konstruktie scoringssysteem van de I.S.I.-Interesstet,
Vorm I en II
Groningen: Wolters
- WESTENDORP, P.H. VAN 1965
Verslag van een onderzoek naar de houding tegenover het parlement,
met enkele methodologische overwegingen betreffende schaalconstructie
Sociologische Gids, 12, 359-373
- WHERRY, R.J. and GAYLORD, M.H. 1943
The concept of test and item reliability in relation to factor
pattern
Psychometrika, 8, 247-264
- WHERRY, R.J. and WINER, B.J. 1953
A method for factoring large numbers of items
Psychometrika, 18, 161-179

WIEGERSMA, S. 1959

Belangstellingsonderzoek bij de differentiatie na de lagere school
Leiden: Batteljie en Terpstra

WIEGERSMA, S. 1960

Psychometrische betrouwbaarheid
Ned. T. Psychol., 15, 22-46

WIEGERSMA, S. en GEMERT, W. 1961

Hoe denkt U over Uw werk?
Mens en Onderneming, 15, 179-192

WOLFF, Ch.J. DE 1963

Personeelsbeoordeling
Amsterdam: Swets en Zeitlinger

WRIGLEY, C.F. 1957

Cluster analysis or factor analysis?
Psychol. Reports, 3, 497-506

I.1. De betrouwbaarheid van de point-biserial r

De betrouwbaarheid van de point-biserial r is afhankelijk van de volgende parameters:

N ; de grootte van de steekproef,

ρ ; de populatie-correlatie,

P_1 ; de proportie één-coderingen van de binaire variabele i ,

P_0 ; de proportie nulcoderingen van de binaire variabele i ,

de verdeling van de scores van de continue variabele,

de verdeling van deze scores met codering één, en

de verdeling van deze scores met codering nul.

Geval 1 : $\rho = 0$

Aannamen:

1. de scores van de continue variabele zijn normaal verdeeld,
2. de scores met codering één zijn normaal verdeeld, en
3. de scores met codering nul zijn normaal verdeeld.

Geval 1a: $\rho = 0$, $N \geq 30$, $0,16 \leq p \leq 0,84$

De hypothese $\rho = 0$ kan bij grote steekproeven worden getoetst met de volgende formule

$$t = r \cdot \sqrt{N}$$

waarbij r de gevonden point-biserial is. Bij $\rho = 0$ heeft t een normale verdeling met gemiddelde nul en sigma één. Deze verdeling staat in nagenoeg alle handboeken, zie onder andere Guilford (1965, tabel B en C), Hoel (1962, tabel II) en McNemar (1963, tabel A).

Geval 1b: $\rho = 0$, $N < 30$, $0,16 \leq p \leq 0,84$

Bij kleine steekproeven kan de hypothese $\rho = 0$ worden getoetst met de volgende formule

$$t' = \frac{r \sqrt{N-2}}{\sqrt{1-r^2}}$$

waarbij r de gevonden point-biserial is. Bij $\rho = 0$ heeft t' een student's t verdeling met $N-2$ vrijheidsgraden. Deze verdeling staat onder andere in Pearson en Hartley (1954, tabel 9).

Geval 2 : $\rho \neq 0$

Aanname :

1. de scores van de continue variabele zijn normaal verdeeld.

Aannamen 2 en 3 van geval 1 zijn, als de populatie correlatie niet nul is, niet acceptabel (Das Gupta, 1960, p. 403 en 404). Om die reden zijn de betrouwbaarheidsgrenzen van Perry en Michael (1954) niet bruikbaar. Deze bezwaren bleken praktisch relevant voor steekproeven kleiner dan 70 van de kapittel-enquête (zie paragraaf 4.1). De correlaties van de steekproeven van 120, 100 en 80 bleken onderling te verschillen overeenkomstig de Fisher z -formule voor produkt moment correlaties. De verschillen met de steekproeven van 60 bleken echter duidelijk groter dan verwacht.

In tabel I.1.1. staan de resultaten van deze toetsing. Deze tabel geeft de eerste twee momenten van de verschillen tussen de z -waarden van de correlaties gedeeld door de theoretisch verwachte standaarddeviatie, waarvan de formule is:

$$\sigma_d = \sqrt{\frac{1}{N_1-3} + \frac{1}{N_2-3}}$$

Tabel I.1.1.

Gemiddelde en standaarddeviatie van de gestandaardiseerde verschillen tussen de z-waarden van de voor 'part-whole' gecorrigeerde point-biserial correlaties van de steekproeven van de kapittel-enquête.

| | N | M | σ |
|-----------------------------|----------|-------|----------|
| populatiewaarden | ∞ | 0 | 1 |
| steekproeven 120, 100 en 80 | 150 | 0,007 | 0,998 |
| steekproeven 60 met andere | 165 | 0,061 | 1,410 |

Gezien deze resultaten kan voorlopig worden aangenomen, dat het praktisch toelaatbaar is om point-biserial's te toetsen met de Fisher z-formule, als de steekproefgrootte groter is dan 70, en de p/q verdeling ligt tussen 0,16 en 0,84.

Geval 2a: $\rho \neq 0$, $N > 70$, $0,16 \leq p \leq 0,84$

Op grond van voorafgaande kan bij grote steekproeven als benaderingsformule in dit geval voor de hypothese $r = \rho$ worden gebruikt:

$$t = |(z\{r\} - z\{\rho\})| \cdot \sqrt{N-3}$$

waarbij $z\{r\}$ de z-waarde van de gevonden point-biserial is, en

$z\{\rho\}$ de z-waarde van de populatie correlatie.

Deze omzetting is mogelijk met de tabel VB van Fisher (1932), tabel H van Guilford (1965) of tabellen B en C van McNemar (1963). De verdeling van t is standaard-normaal.

De hypothese $r_1 = r_2$ wordt dan getoetst met de formule:

$$t = (z\{r_1\} - z\{r_2\}) \cdot \sqrt{\frac{(N_1-3)(N_2-3)}{N_1+N_2-6}}$$

Ook hier is de verdeling van t standaard-normaal. Een belangrijke voorwaarde van de toetsing is, dat r_1 en r_2 zijn uitgerekend in twee onafhankelijke 'unmatched' steekproeven.

Als niet aan bovengenoemde voorwaarde is voldaan kan de hypothese $r_{12} = r_{13}$ worden getoetst met de z.g. Hotelling-formule (1940), zie Guilford (1965, formule 9.12):

$$t = (r_{12} - r_{13}) \cdot \sqrt{\frac{(N-3)(1+r_{23})}{2(1-r_{23}^2 - r_{12}^2 - r_{13}^2 + 2r_{23}r_{12}r_{13})}}$$

Hierbij dient r_{23} een produkt moment correlatie te zijn. Deze toetsing is o.a. bruikbaar als moet worden nagegaan of de ladingen van een operator in twee clusters significant verschillen. De clusters moeten daarbij minstens ieder uit 15 operatoren bestaan, anders dient een correctie te worden aangebracht voor 'part-whole' correlatie (zie paragraaf 6.3.1.). Bij clusters kleiner dan 10 daarnaast ook een correctie voor 'coarse-grouping' (zie paragraaf 6.3.1.). De t -waarde van de Hotelling-formule is ook standaard-normaal verdeeld.

Geval 2b : $\rho \neq 0$, $N \leq 70$

Er zijn geen formules, ook geen benaderingsformules om in dit geval de point-biserial te toetsen.

1.2. Toetsing van een reeks afhankelijke r-waarden.

Deze toetsing is mogelijk voor de operatoren van één cluster bij een steekproefgrootte van $N \geq 200$ en wel op de volgende wijze: rangschik de correlaties van hoog naar laag en begin met de toetsing van het grootste verschil, namelijk van de grootste en kleinste correlatie en wel met de Hotelling formule van geval 2a (zie bijlage I.1.). De intercorrelatie tussen de twee operatoren moet in dit geval een tetrachorische r zijn. Deze is snel te berekenen met de tabellen van Jenkins (1955), ook voor p/q verdelingen, die afwijken van de 50/50%. Op grond van het feit dat de tetrachorische r pas bij steekproeven groter dan 200 nauwkeurige schattingen van de produkt moment correlatie geeft (Guilford, 1965, p. 330 en 331), moet deze toetsing bij kleinere steekproeven beschouwd worden als een benadering.

De significantie-drempel is bij bovengenoemde toetsing gelijk aan (Ryan, 1960)

$$\alpha_1 = \frac{2}{n(K-1)}$$

Hierbij is α de normale significantiegrens (5% of 1%),

n het totaal aantal operatoren, en

K het aantal operatoren van de sub-steekproef.

Bij bovengenoemde toetsing is bij de eerste toetsing K gelijk aan n . Wanneer de twee correlaties significant verschillen, wordt op dezelfde wijze het op één na grootste verschil getoetst. K is dan echter niet n , maar $n-1$. Met deze toetsingen wordt doorgegaan, waarbij K steeds met één wordt verminderd, totdat het verschil niet meer significant is. Dit is de zogenaamde 'plugging in' methode met aangepaste grenzen. De

correlaties van de operatoren van de sub-steekproef, waar voor het eerst het grootste verschil niet-significant is, zijn homogeen. Bij de correlaties kan dan de nulhypothese:

$$r_1 = r_2 = \dots = r_i = \dots = r_K = \rho$$

niet worden verworpen.

De beste schatting van de populatie correlatie is dan

$$\hat{\rho} = r \left\{ \sum_i^K \frac{z\{r_i\}}{K} \right\}$$

Voor de omzetting van de correlatiecoëfficiënten in Fisher z-waarden kunnen de tabellen van Fisher (1932, tabel VB), van Guilford (1965, tabel H) of van McNemar (1963, tabel B en C) worden gebruikt.

I.3. Toetsing van een reeks afhankelijke p-waarden.

McNemar (1963, formule 13.5) heeft een formule ontwikkeld om een reeks van afhankelijke proporties op hun homogeniteit te toetsen. Deze formule is bruikbaar bij steekproeven waarvan $N > 30$ en luidt

$$Q = \frac{(n-1) \left[n \sum_i (Np_i)^2 - \left(\sum_i Np_i \right)^2 \right]}{n \sum_b S_b^2 - \sum_b S_b^2}$$

Hierbij is n het aantal operatoren ($1 \dots i \dots n$),

p_i de proportie één-coderingen van operator i ,

N het aantal proefpersonen of stimuli ($1 \dots b \dots N$), en

S_b de somscore van proefpersoon b over alle n operatoren.

De waarde Q volgt een chi-kwadraat verdeling met $n-1$ vrijheidsgraden.

Deze is voor meer dan 30 vrijheidsgraden te berekenen met de volgende formule:

$$\chi^2 \{P \leq \alpha\} = n + 0,5 \cdot t^2 \{P \leq \alpha\} + t \{P \leq \alpha\} \cdot \sqrt{2(n-1)} - 1,5$$

Hierbij is α de significantiegrens,

t de t -waarde van de standaard-normaalverdeling voor de gekozen significantiegrens.

De standaard-normaalverdeling staat in nagenoeg alle handboeken, onder andere in Guilford (1965, tabel B en C), McNemar (1963, tabel A) en Hoel (1962, tabel II).

Voor minder dan 30 vrijheidsgraden kunnen de tabellen van Fisher (1932, tabel III), Guilford (1965, tabel E) en McNemar (1962, tabel D) worden gebruikt.

I.4. Betrouwbaarheid van de K.R.(21) en K.R.(20).

Hoewel er al verschillende pogingen zijn ondernomen om de verdeling van de waarden van de K.R. formules in steekproeven te bepalen, moet toch nog gezegd worden, dat de resultaten niet definitief en afgerond zijn.

Lord (1955) heeft een onderscheid gemaakt tussen drie vormen van steekproeftrekking. Bij het eerste type worden dezelfde operatoren toegepast op aselekt getrokken steekproeven van elementen. Het gaat dan om de statistische betrouwbaarheid. Wanneer nu aselechte sets van operatoren worden toegepast op dezelfde steekproef van elementen, kan worden gesproken over type twee 'sampling of randomly parallel forms or tests'. Het onderzoek richt zich dan op de psychometrische betrouwbaarheid. Combinatie van beide types geeft de derde vorm; aselechte sets van operatoren bij aselechte steekproeven van elementen.

Voor het tweede geval geeft Lord (1955) zelf betrouwbaarheidsformules, voor het derde geval Feldt (1965), terwijl naar onze mening Leyrly (1959) een toets heeft gegeven voor de hypothese K.R.(21) is nul in geval 1. Vrij recent hebben Payne en Anderson (1968) door middel van simulatie significantie niveau's bepaald voor de K.R.(20). De door hen gevolgde procedure is echter aanvechtbaar. Denkelijk wordt met hun tabel alleen getoetst of in het door de onderzoeker gevonden cluster de invloed van de meetfout op de score significant kleiner is dan in de door Payne en Anderson gesimuleerde data. Om die reden achten wij de tabellen van Payne en Anderson niet bruikbaar.

De toetsingen van de drie gevallen zijn nu als volgt:

Geval 1:

De vraag luidt: hoe groot is de kans dat dit cluster bij een steekproef van elementen uit dezelfde populatie weer wordt gevonden?

Voor de evaluatie van gevonden clusters is deze vraag in de eerste plaats belangrijk. Leyrly (1954) heeft erop gewezen dat in dit geval de hypothese $K.R.(21) = 0$ getoetst kan worden met de binomiale index van dispersie (zie Hoel, 1962, p. 255-257). Nagegaan wordt of de verdeling van de clusterscores binomiaal is. Als dat het geval is, is de kans klein dat de cluster bij een steekproef van overeenkomstige elementen weer wordt gevonden. Een binomiale verdeling van de scores ontstaat namelijk bij onafhankelijke operatoren en/of elementen met gelijke ware scores. In beide gevallen meet de schaal of cluster in deze populatie van elementen niets.

De formule luidt:

$$\chi^2 = \frac{n \cdot N}{n - (n-1) \cdot r_{21}}$$

Hierbij is χ^2 de chi-kwadraat van deze $K.R.(21)$, bij deze N en n ,
 n het aantal operatoren,
 N de grootte van de steekproef van elementen, en
 r_{21} de berekende $K.R.(21)$

De chi-kwadraat verdeelt zich met $(N-1)$ vrijheidsgraden. In Fisher (1932, tabel III), Guilford (1965, tabel E), McNemar (1962, tabel D) staat deze verdeling voor 1 t/m 30 vrijheidsgraden. Bij meer dan dertig vrijheidsgraden kan de standaard-normaalverdeling worden gebruikt voor de t -waarde, waarbij

$$t = \sqrt{2\chi^2} - \sqrt{2N-3}$$

Het is ook mogelijk de K.R.(21) te berekenen, die bij deze N en n significant is voor bijv. 5 of 1%, en wel met de formule:

$$r_{21} \{ \alpha \geq P \} = \frac{n}{n-1} \left(1 - \frac{N}{\chi^2 \{ \alpha \geq P \}} \right)$$

waarbij voor meer dan dertig vrijheidsgraden

$$\chi^2 \{ \alpha \geq P \} = \frac{1}{2} \cdot t^2 \{ \alpha \geq P \} + \sqrt{2N-3} \cdot t \{ \alpha \geq P \} + N - \frac{1}{2}$$

Voor de K.R.(20) heeft van der Ven gewezen op de index van dispersie voor de gegeneraliseerde binomiaal. In dat geval is de formule:

$$\chi^2 = \frac{N \cdot \sigma^2}{\sum_i p_i \left\{ 1 - \frac{\sum p_i}{n} \right\}}$$

Wanneer het aantal operatoren meer is dan twintig kan de betrouwbaarheid van de K.R.(20) bepaald worden met de formules van geval 3. Lord (1955) heeft aangetoond, dat de meetfout van geval 3 gelijk is aan de som van de meetfouten van geval 1 en 2. Bij veel operatoren is de meetfout van geval 2 verwaarloosbaar klein en kan dus de meetfout van geval 3 gelijk gesteld worden aan geval 1 (zie ook Feld, 1965, p. 364). Hierbij dient wel opgemerkt te worden dat de grens van twintig arbitrair is en nog niet door onderzoek geverifieerd.

Geval 2:

De vraag is: hoe groot is de kans dat een overeenkomstig cluster bij dezelfde elementen weer wordt gevonden? Deze vraag is o.a. belangrijk bij kennistoetsingen, wanneer de onderzoeker wil weten hoe betrouwbaar de klassificatie voldoende-onvoldoende is op grond van een aantal opdrachten of vragen.

Lord (1955) heeft voor dit geval formules ontwikkeld. Deze formules, die wij hier niet zullen geven, hebben echter alleen een theoretische waarde. Ze zijn namelijk alleen toepasbaar bij clusters van redelijke grootte. In dat geval is echter de meetfout van geval 2 verwaarloosbaar klein (Lord, 1955, p. 6 en 19).

Geval 3:

De vraag is: hoe groot is de kans dat een overeenkomstig cluster bij een steekproef van elementen uit de populatie weer wordt gevonden?

Deze vraag is o.a. belangrijk voor examens waarbij steeds andere vragen uit eenzelfde stock aan verschillende kandidaten worden voorgelegd en ook bij het gebruik van parallel tests bij verschillende groepen uit de zelfde populatie.

Feldt (1965) geeft een benadering via variantie-analyse van de betrouwbaarheid van de K.R.(20) in dit geval. De formule luidt:

$$P \left\{ 1 - \frac{(1-\rho_{20})}{F_a} < r_{20} < 1 - \frac{(1-\rho_{20})}{F_b} \right\} = 1 - \alpha$$

Hierbij is α de gekozen significantiegrens,

ρ_{20} de populatie of gestelde K.R.(20),

r_{20} de verkregen K.R.(20),

F_a het linker percentiel punt in de centrale F-verdeling voor $\frac{1}{2}\alpha$

F_b het rechter percentiel punt in de centrale F-verdeling voor $1 - \frac{1}{2}\alpha$

Het aantal vrijheidsgraden voor F_a en F_b zijn $(N-1)$ en $(N-1)(n-1)$. De eerste voor de teller (ordinaat van de tabel), de tweede voor de noemer (absis van de tabel). Daar in de meeste tabellen alleen de rechter overschrijdingskans wordt gegeven (F_b), kan F_a berekend worden met de relatie:

$$F \{P \leq \frac{1}{2}\alpha, df_1, df_2\} = \frac{1}{F \{P \leq 1-\frac{1}{2}\alpha, df_2, df_1\}}$$

Hierbij moet erop worden gewezen dat de vrijheidsgraden worden omgedraaid; df_2 wordt de teller en df_1 de noemer. Lindquist (1953) en Haggard (1958) geven bovenstaande formule verkeerd doordat ze dit punt over het hoofd zien. Ook Ebel (1951) geeft de formule echter om andere redenen onjuist. De formule voor de betrouwbaarheidsinterval van ρ_{20} wordt:

$$P \{1 - (1-r_{20}) \cdot F_b < \rho_{20} < 1 - (1-r_{20}) F_a\} = 1 - \alpha$$

Tabellen voor de centrale F-verdeling staan in Snedecor (1937), Hoel (1962, tabel III), McNemar (1963, tabel F) en Guilford (1965, tabel F).

Bovenstaande formules moeten als benaderingsformules worden beschouwd.

Bij nul en één-coderingen is niet aan de voorwaarden voldaan die gesteld worden voor toetsingen door middel van variantie-analyse. In een empirisch onderzoek vindt Feldt echter geen grote verschillen.

I.5. Omzetting van de point-biserial r in de biserial r.

In tabel I.5.1. staan de coëfficiënten waarmee de gevonden point-biserial r moet worden vermenigvuldigd om een biserial r te verkrijgen. De coëfficiënt is afhankelijk van de p- of q-waarde van de variabele; de proportie enen of nullen.

Tabel I.5.1.

Coëfficiënten voor omzetting
point-biserial r in biserial r

| p of q | coëff. | p of q | coëff. | p of q | coëff. | p of q | coëff. |
|--------|--------|--------|--------|--------|--------|--------|--------|
| | | 0,15 | 1,53 | 0,27 | 1,34 | 0,39 | 1,27 |
| | | 0,16 | 1,51 | 0,28 | 1,33 | 0,40 | 1,27 |
| 0,05 | 2,11 | 0,17 | 1,48 | 0,29 | 1,33 | 0,41 | 1,27 |
| 0,06 | 1,99 | 0,18 | 1,46 | 0,30 | 1,32 | 0,42 | 1,26 |
| 0,07 | 1,90 | 0,19 | 1,45 | 0,31 | 1,31 | 0,43 | 1,26 |
| 0,08 | 1,83 | 0,20 | 1,43 | 0,32 | 1,30 | 0,44 | 1,26 |
| 0,09 | 1,76 | 0,21 | 1,41 | 0,33 | 1,30 | 0,45 | 1,26 |
| 0,10 | 1,71 | 0,22 | 1,40 | 0,34 | 1,29 | 0,46 | 1,25 |
| 0,11 | 1,66 | 0,23 | 1,39 | 0,35 | 1,29 | 0,47 | 1,25 |
| 0,12 | 1,62 | 0,24 | 1,37 | 0,36 | 1,28 | 0,48 | 1,25 |
| 0,13 | 1,59 | 0,25 | 1,36 | 0,37 | 1,28 | 0,49 | 1,25 |
| 0,14 | 1,56 | 0,26 | 1,35 | 0,38 | 1,27 | 0,50 | 1,25 |

BIJLAGE II.1. DATASIMULATIE ONDERZOEKEN (ZIE PARAGRAAF 6.1)

DATASIMULATIE I
ASPEKTE DATA VAN 20 ELEMENTEN BIJ 50 OPERATIEËN

| | | | | | | | | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1 | 49 | 58 | 52 | 55 | 56 | 26 | 47 | 53 | 66 | 49 | 57 | 41 | 52 | 47 | 49 | 64 | 54 | 42 | 41 |
| | 51 | 49 | 48 | 45 | 52 | 51 | 71 | 63 | 48 | 52 | 48 | 38 | 53 | 59 | 50 | 59 | 53 | 27 | 47 |
| | 42 | 54 | 57 | 42 | 58 | 57 | 59 | 72 | 49 | 53 | 56 | 53 | | | | | | | |
| 2 | 33 | 53 | 43 | 54 | 62 | 64 | 57 | 56 | 57 | 42 | 56 | 52 | 63 | 36 | 60 | 56 | 32 | 26 | 51 |
| | 54 | 56 | 45 | 44 | 46 | 52 | 56 | 35 | 60 | 54 | 26 | 51 | 50 | 44 | 34 | 63 | 50 | 50 | 58 |
| | 61 | 64 | 56 | 41 | 50 | 41 | 66 | 57 | 65 | 51 | 46 | 60 | | | | | | | |
| 3 | 55 | 54 | 55 | 49 | 62 | 55 | 43 | 32 | 45 | 57 | 41 | 41 | 54 | 60 | 53 | 58 | 53 | 50 | 65 |
| | 39 | 59 | 40 | 45 | 30 | 52 | 52 | 52 | 30 | 50 | 30 | 37 | 60 | 47 | 51 | 64 | 42 | 46 | 44 |
| | 66 | 50 | 54 | 38 | 23 | 30 | 38 | 45 | 41 | 50 | 48 | 51 | | | | | | | |
| 4 | 43 | 41 | 46 | 49 | 50 | 53 | 55 | 59 | 36 | 62 | 50 | 49 | 49 | 61 | 39 | 51 | 46 | 69 | 71 |
| | 59 | 48 | 65 | 47 | 47 | 45 | 49 | 50 | 43 | 53 | 46 | 54 | 46 | 57 | 49 | 56 | 36 | 54 | 56 |
| | 44 | 66 | 54 | 46 | 50 | 52 | 47 | 58 | 34 | 48 | 38 | 45 | | | | | | | |
| 5 | 42 | 34 | 41 | 52 | 59 | 50 | 54 | 64 | 63 | 37 | 54 | 50 | 45 | 51 | 53 | 49 | 74 | 56 | 63 |
| | 54 | 55 | 70 | 59 | 57 | 44 | 54 | 57 | 45 | 27 | 37 | 25 | 35 | 54 | 45 | 63 | 56 | 66 | 41 |
| | 53 | 49 | 45 | 34 | 69 | 57 | 36 | 41 | 46 | 55 | 36 | 34 | | | | | | | |
| 6 | 53 | 61 | 53 | 55 | 28 | 61 | 48 | 50 | 43 | 48 | 53 | 44 | 61 | 54 | 42 | 46 | 36 | 52 | 31 |
| | 46 | 57 | 29 | 48 | 56 | 62 | 42 | 62 | 51 | 41 | 43 | 56 | 35 | 57 | 62 | 42 | 47 | 45 | 32 |
| | 47 | 56 | 40 | 42 | 47 | 43 | 61 | 56 | 51 | 43 | 68 | 45 | | | | | | | |
| 7 | 60 | 40 | 33 | 49 | 48 | 61 | 49 | 41 | 49 | 45 | 53 | 47 | 40 | 44 | 52 | 53 | 48 | 49 | 55 |
| | 48 | 35 | 56 | 63 | 52 | 43 | 50 | 40 | 48 | 36 | 40 | 42 | 40 | 49 | 57 | 44 | 64 | 31 | 52 |
| | 55 | 58 | 51 | 50 | 47 | 63 | 52 | 56 | 71 | 59 | 66 | 37 | | | | | | | |
| 8 | 53 | 40 | 59 | 55 | 38 | 53 | 47 | 39 | 55 | 54 | 41 | 53 | 63 | 52 | 42 | 53 | 55 | 40 | 51 |
| | 56 | 55 | 39 | 54 | 39 | 47 | 35 | 58 | 56 | 51 | 44 | 50 | 47 | 54 | 65 | 62 | 38 | 50 | 29 |
| | 55 | 45 | 58 | 58 | 53 | 47 | 57 | 37 | 63 | 35 | 52 | 60 | | | | | | | |
| 9 | 43 | 45 | 45 | 38 | 42 | 59 | 50 | 48 | 32 | 59 | 45 | 54 | 66 | 55 | 53 | 58 | 49 | 61 | 61 |
| | 27 | 60 | 54 | 56 | 47 | 56 | 35 | 48 | 38 | 49 | 51 | 40 | 51 | 53 | 57 | 45 | 50 | 50 | 62 |
| | 59 | 48 | 59 | 50 | 60 | 53 | 48 | 50 | 66 | 43 | 57 | 48 | | | | | | | |
| 10 | 69 | 50 | 47 | 60 | 43 | 44 | 45 | 55 | 42 | 52 | 54 | 53 | 54 | 47 | 54 | 45 | 46 | 60 | 24 |
| | 67 | 61 | 39 | 55 | 64 | 49 | 57 | 52 | 53 | 67 | 57 | 59 | 24 | 49 | 44 | 53 | 38 | 45 | 56 |
| | 44 | 40 | 44 | 32 | 63 | 43 | 58 | 44 | 33 | 35 | 61 | 54 | | | | | | | |
| 11 | 58 | 56 | 39 | 37 | 45 | 36 | 49 | 40 | 52 | 47 | 41 | 52 | 60 | 44 | 45 | 33 | 42 | 44 | 26 |
| | 46 | 70 | 45 | 48 | 55 | 51 | 37 | 48 | 55 | 54 | 32 | 45 | 66 | 47 | 45 | 49 | 66 | 43 | 50 |
| | 44 | 48 | 44 | 27 | 53 | 76 | 67 | 44 | 62 | 34 | 48 | 62 | | | | | | | |
| 12 | 51 | 50 | 57 | 46 | 59 | 44 | 54 | 60 | 68 | 37 | 52 | 51 | 42 | 41 | 57 | 59 | 46 | 72 | 24 |
| | 45 | 38 | 46 | 63 | 48 | 48 | 42 | 53 | 37 | 56 | 50 | 27 | 48 | 36 | 54 | 59 | 57 | 63 | 56 |
| | 45 | 56 | 64 | 44 | 63 | 46 | 48 | 55 | 42 | 52 | 35 | 69 | | | | | | | |

DATASIMULATIE I
 ASELEKTE DATA VAN 20 ELEMENTEN BIJ 50 OPERATOREN

VERVOLG

| | | | | | | | | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 13 | 54 | 69 | 48 | 58 | 53 | 56 | 61 | 46 | 44 | 40 | 51 | 57 | 64 | 55 | 46 | 41 | 63 | 64 | 65 |
| | 25 | 40 | 41 | 61 | 31 | 60 | 53 | 48 | 42 | 77 | 41 | 61 | 52 | 68 | 33 | 58 | 70 | 40 | 54 |
| | 58 | 62 | 65 | 48 | 72 | 50 | 34 | 31 | 54 | 57 | 52 | 41 | | | | | | | |
| | | | | | | | | | | | | | | | | | | | |
| 14 | 49 | 46 | 65 | 59 | 54 | 53 | 40 | 47 | 57 | 47 | 48 | 49 | 59 | 57 | 49 | 50 | 42 | 56 | 68 |
| | 49 | 72 | 46 | 50 | 44 | 35 | 54 | 62 | 58 | 57 | 47 | 53 | 44 | 61 | 44 | 54 | 46 | 47 | 45 |
| | 57 | 53 | 59 | 45 | 59 | 22 | 43 | 59 | 55 | 45 | 37 | 52 | | | | | | | |
| | | | | | | | | | | | | | | | | | | | |
| 15 | 83 | 45 | 60 | 63 | 45 | 45 | 43 | 54 | 52 | 53 | 60 | 16 | 39 | 37 | 45 | 43 | 51 | 57 | 66 |
| | 43 | 57 | 45 | 50 | 22 | 45 | 55 | 50 | 31 | 29 | 63 | 57 | 52 | 30 | 57 | 56 | 48 | 45 | 45 |
| | 53 | 54 | 61 | 61 | 47 | 69 | 36 | 56 | 47 | 51 | 53 | 56 | | | | | | | |
| | | | | | | | | | | | | | | | | | | | |
| 16 | 45 | 41 | 63 | 59 | 35 | 53 | 52 | 73 | 43 | 51 | 36 | 59 | 57 | 52 | 43 | 60 | 45 | 55 | 65 |
| | 70 | 50 | 60 | 42 | 45 | 33 | 61 | 50 | 55 | 61 | 60 | 44 | 39 | 60 | 54 | 42 | 50 | 59 | 41 |
| | 44 | 61 | 49 | 61 | 52 | 48 | 42 | 46 | 68 | 59 | 53 | 30 | | | | | | | |
| | | | | | | | | | | | | | | | | | | | |
| 17 | 49 | 53 | 57 | 57 | 61 | 53 | 64 | 46 | 44 | 50 | 43 | 56 | 59 | 41 | 56 | 37 | 48 | 63 | 57 |
| | 43 | 36 | 54 | 41 | 48 | 41 | 45 | 47 | 60 | 59 | 49 | 53 | 58 | 53 | 41 | 51 | 50 | 43 | 47 |
| | 54 | 51 | 64 | 35 | 37 | 44 | 53 | 46 | 67 | 65 | 56 | 52 | | | | | | | |
| | | | | | | | | | | | | | | | | | | | |
| 18 | 53 | 43 | 64 | 46 | 35 | 47 | 48 | 55 | 57 | 49 | 42 | 46 | 52 | 23 | 41 | 38 | 56 | 44 | 53 |
| | 54 | 39 | 50 | 46 | 46 | 49 | 49 | 41 | 40 | 59 | 54 | 42 | 53 | 63 | 66 | 52 | 44 | 46 | 52 |
| | 59 | 53 | 57 | 37 | 32 | 33 | 63 | 41 | 40 | 40 | 57 | 63 | | | | | | | |
| | | | | | | | | | | | | | | | | | | | |
| 19 | 35 | 48 | 62 | 47 | 43 | 48 | 43 | 46 | 38 | 68 | 47 | 47 | 49 | 61 | 42 | 48 | 46 | 51 | 50 |
| | 58 | 49 | 54 | 42 | 50 | 33 | 47 | 53 | 47 | 63 | 50 | 60 | 34 | 36 | 55 | 51 | 39 | 71 | 44 |
| | 75 | 51 | 60 | 54 | 45 | 51 | 34 | 40 | 45 | 53 | 39 | 63 | | | | | | | |
| | | | | | | | | | | | | | | | | | | | |
| 20 | 59 | 61 | 42 | 43 | 57 | 72 | 41 | 72 | 56 | 65 | 36 | 75 | 59 | 50 | 49 | 59 | 69 | 59 | 45 |
| | 52 | 50 | 57 | 50 | 54 | 39 | 50 | 31 | 47 | 36 | 37 | 58 | 35 | 41 | 55 | 34 | 34 | 57 | 49 |
| | 50 | 43 | 57 | 61 | 73 | 59 | 59 | 47 | 30 | 45 | 51 | 33 | | | | | | | |

DATASIMULATIE 11
 ASELECTE DATA VAN 100 ELEMENTEN BIJ 50 OPERATOREN

```

1  1  20  44  34  63  50  50  58  61  64  56  41  50  41  66  57  65  51  46
   61  53  54  55  49  62  55  43  32  45  57  41  41  54  60  53  58  53  50
   55  59  59  47  45  30  52  52  52  30  50  30
2  37  60  47  51  64  42  46  44  66  50  54  38  23  51  38  45  41  50  48
   51  45  41  40  49  50  53  55  59  36  62  50  49  49  61  39  51  46  69
   71  59  46  65  47  47  45  49  50  43  53  46
3  54  46  57  49  56  36  54  56  44  66  54  46  50  52  47  56  34  48  38
   45  42  34  41  52  59  50  54  64  63  37  54  50  45  51  53  49  74  56
   55  54  55  70  59  57  44  54  57  45  27  37
4  25  55  54  45  63  56  66  41  53  49  45  34  69  57  36  41  46  55  36
   54  53  61  53  55  28  61  48  50  45  48  53  44  61  54  42  46  36  52
   51  46  57  29  48  56  62  42  62  51  41  43
5  55  55  57  62  42  47  45  33  47  56  40  42  47  43  61  56  51  43  68
   45  66  40  35  49  48  61  49  41  49  45  53  47  40  44  52  53  48  49
   55  48  35  56  63  52  43  50  40  48  36  40
6  52  40  49  57  44  64  31  52  55  58  51  50  47  63  52  56  71  59  66
   57  53  40  59  55  28  53  47  39  55  54  41  53  63  52  42  53  55  40
   51  56  55  59  54  39  47  35  58  56  51  44
7  50  47  54  65  62  38  59  29  55  45  58  58  53  47  57  37  63  33  52
   60  43  45  43  38  42  59  50  48  32  59  45  54  56  53  53  58  49  61
   61  27  60  54  56  47  56  35  48  36  49  51
8  40  51  53  57  45  50  50  63  59  48  59  50  60  53  48  50  56  43  57
   46  69  50  47  60  43  44  45  55  42  52  54  53  54  47  54  45  46  60
   24  57  61  39  55  64  49  57  52  53  67  57
9  59  24  49  44  53  38  45  56  44  40  44  32  63  43  58  44  33  35  61
   54  58  56  39  37  45  36  49  40  52  47  41  52  60  44  45  33  42  44
   26  46  70  45  48  55  51  37  48  55  54  32
10  45  66  47  45  49  66  43  50  44  48  44  27  53  76  67  44  62  34  48
   62  51  50  57  46  59  44  54  60  68  37  52  51  42  41  57  59  46  72
   54  43  38  46  63  48  48  42  53  37  56  50
11  27  48  56  54  59  57  63  56  45  56  64  44  63  46  48  55  42  52  35
   69  54  69  48  58  53  56  61  46  44  40  51  57  64  55  46  41  63  64
   55  25  46  41  61  31  60  53  48  42  77  41
12  51  52  58  33  58  70  40  54  58  62  65  48  72  50  34  31  54  57  52
   41  49  46  65  59  54  55  40  47  57  47  48  49  59  57  49  50  42  56
   58  49  72  46  50  44  35  54  62  58  57  47
13  53  44  51  44  54  46  47  45  57  53  59  45  59  22  43  59  55  45  37
   52  33  45  60  63  45  45  43  54  52  53  60  16  39  37  45  43  51  57
   66  43  57  45  50  22  45  55  50  31  29  63
14  57  52  50  57  56  48  45  45  53  54  61  61  47  69  36  56  47  51  53
   56  45  41  63  59  35  53  52  73  43  51  36  59  57  52  43  60  45  55
   55  70  50  60  42  45  33  61  50  55  61  60
15  44  39  60  54  42  50  59  41  44  61  49  61  52  48  42  46  68  59  53
   30  49  53  57  57  61  53  64  46  44  50  43  56  59  41  56  37  48  63
   57  43  36  54  41  48  41  45  47  60  59  49
16  53  58  53  41  51  50  43  47  54  51  64  35  37  44  53  46  67  65  56
   52  53  43  64  46  35  47  48  55  57  49  42  46  52  23  41  38  56  44
   53  54  39  50  46  46  49  49  41  40  59  54

```

DATASIMULATIE II
 ASELEKTE DATA VAN 100 ELEMENTEN BIJ 50 OPERATOREN

VERVULG

17 +2 53 53 66 52 44 46 52 59 53 57 37 32 33 63 41 40 40 57
 63 35 40 62 47 43 48 43 46 38 68 47 47 49 61 42 48 46 51
 50 58 49 54 42 50 33 47 53 47 63 50
 18 60 34 36 55 51 39 71 44 75 51 60 54 45 51 54 40 45 53 39
 63 59 61 42 43 57 72 41 72 56 65 36 75 59 50 49 59 69 50
 45 52 50 57 50 54 39 50 31 47 36 37
 19 53 35 41 55 34 34 57 49 50 43 57 61 73 59 59 47 30 45 51
 35 60 64 32 51 34 62 44 40 65 45 41 43 48 42 50 26 49 34
 47 45 36 55 32 70 31 57 46 39 44 59
 20 52 57 52 48 56 56 48 54 64 54 42 50 66 58 53 39 51 28 42
 57 42 46 64 63 44 66 43 35 44 53 50 45 38 64 44 46 38 40
 36 54 45 59 76 46 53 46 48 50 48 57
 21 39 48 57 57 48 52 40 55 44 57 58 42 71 58 74 40 62 46 42
 44 63 49 62 59 52 60 40 53 48 52 58 67 34 52 47 42 43 18
 54 62 61 38 43 42 73 34 53 40 41 51
 22 55 45 43 62 59 55 45 45 40 73 69 48 54 45 53 34 58 57 46
 54 51 41 47 33 45 56 59 38 53 49 46 60 39 59 63 48 52 64
 70 46 39 61 65 55 72 51 41 49 63 66
 23 60 57 36 39 48 48 38 55 52 53 46 42 61 57 61 50 57 53 46
 49 63 40 42 53 52 26 47 36 61 52 48 45 60 40 34 48 58 56
 47 58 58 50 49 34 43 48 46 23 51 52
 24 45 50 57 50 54 43 44 47 44 31 41 48 48 58 42 40 61 56 33
 44 47 61 42 61 46 60 56 55 53 53 58 49 55 60 49 42 30 56
 56 53 48 54 43 52 49 51 51 51 60 46
 25 58 27 54 50 69 44 63 55 35 52 43 55 54 40 30 56 74 51 44
 47 53 39 54 37 45 59 44 58 41 49 38 44 48 48 52 60 73 39
 54 74 55 35 55 44 65 50 50 52 55 42
 26 52 52 31 39 57 47 49 63 40 35 58 27 59 55 48 54 56 56 60
 32 45 61 51 38 70 45 54 53 28 54 49 49 49 45 44 50 42 40
 62 41 50 77 46 51 48 43 42 69 32 52
 27 56 47 70 43 57 47 42 43 58 35 61 45 30 38 49 71 68 65 32
 47 49 47 40 48 51 67 50 54 46 58 44 45 34 39 34 53 30 56
 47 44 50 70 59 32 53 34 40 66 57 50
 28 57 46 54 36 56 54 38 38 61 43 34 67 57 38 60 53 33 48 54
 71 50 51 57 44 58 39 53 41 32 34 46 41 37 42 63 43 41 30
 61 51 44 34 52 60 56 47 58 45 51 41
 29 53 46 44 40 33 52 52 41 46 36 34 60 51 61 29 59 69 63 42
 39 36 60 56 53 45 52 37 57 35 60 44 49 74 47 37 47 53 38
 47 53 55 52 29 50 75 56 49 53 40 71
 30 42 48 47 60 51 35 58 48 34 53 55 51 50 55 61 48 62 65 52
 48 38 50 47 44 34 38 46 40 30 58 39 42 48 52 40 46 34 74
 56 54 42 59 52 65 33 45 48 55 44 61
 31 25 57 40 78 50 35 72 49 36 70 48 50 51 51 44 36 50 44 50
 50 63 75 37 51 22 37 58 44 62 62 49 59 48 39 47 61 48 40
 53 59 52 59 68 52 42 52 64 44 61 37
 32 44 36 55 43 53 45 40 61 29 47 66 55 52 43 34 50 44 44 46
 53 50 52 48 64 64 62 66 52 45 61 59 51 51 48 36 56 45 63
 51 42 48 49 56 48 57 60 40 54 48 62

DATASIMULATIE II
 ASELKTE DATA VAN 100 ELEMENTEN BIJ 50 OPERATIEEN

VERVOLG

```

33 37 43 50 43 61 48 53 53 59 55 41 33 48 52 47 48 53 48 28
    51 53 55 62 65 48 37 40 54 40 47 52 60 55 25 43 47 42 37
    54 50 66 58 47 55 44 44 46 41 39 45
34 50 42 41 46 56 40 25 46 65 50 48 54 52 44 46 43 45 70 49
    56 56 49 50 50 57 51 42 52 53 42 59 51 49 65 40 60 54 59
    44 61 57 43 50 55 72 38 23 53 51 46
35 67 60 51 64 48 32 46 37 46 41 43 35 60 71 56 57 52 50 61
    47 58 56 33 58 42 48 38 56 43 50 46 61 60 54 47 61 45 77
    54 58 42 46 41 56 65 71 54 40 63 63
36 52 58 43 57 55 50 48 36 45 56 40 48 54 44 63 62 50 44 45
    42 40 48 44 62 42 69 40 35 59 59 55 54 54 49 54 58 51 54
    53 57 57 70 59 43 52 45 42 41 45 58
37 44 55 31 42 38 44 42 40 50 58 55 59 56 47 59 54 37 65 43
    45 42 57 60 54 45 40 43 54 51 44 37 49 36 59 42 56 48 28
    35 54 39 55 68 64 70 57 34 57 45 52
38 69 48 60 30 59 44 44 44 53 62 45 68 51 41 46 56 58 65 60
    43 45 46 33 21 71 40 51 44 43 55 54 35 71 59 42 22 38 44
    51 50 50 53 50 46 52 40 40 41 44 43
39 69 48 59 70 47 42 61 42 45 29 42 51 39 58 57 51 54 48 42
    49 48 62 49 52 60 74 45 50 53 69 45 46 49 55 71 62 65 42
    52 55 47 32 55 31 55 47 60 51 69 62
40 58 44 33 46 55 54 30 56 35 57 49 55 56 58 46 63 57 47 53
    37 40 44 62 31 67 51 47 36 45 52 64 60 80 54 60 57 60 55
    36 49 56 45 53 33 62 55 35 57 40 49
41 46 50 46 63 45 53 54 44 60 51 38 61 62 67 52 35 65 59 40
    40 62 67 35 50 51 40 38 62 50 62 47 54 59 68 53 58 51 62
    61 46 49 46 54 31 60 37 59 61 62 59
42 56 72 54 34 55 46 57 50 49 53 46 54 62 47 56 35 50 44 44
    45 31 42 49 57 47 49 62 56 47 57 56 60 37 48 55 47 52 48
    57 48 50 53 60 56 49 42 53 46 62 36
43 50 64 41 47 65 58 52 39 53 48 41 48 32 67 57 52 42 49 52
    52 45 59 46 28 45 45 56 56 35 37 46 59 55 35 56 56 52 44
    55 55 28 29 66 49 36 47 51 60 80 36
44 38 32 37 38 42 63 36 35 35 61 37 41 46 47 51 44 60 64 55
    52 61 64 42 44 66 46 61 43 54 65 53 63 46 44 37 50 54 62
    49 60 42 45 53 55 58 57 36 45 38 45
45 38 44 44 50 40 47 47 52 54 51 50 54 26 65 71 52 51 49 45
    41 49 47 25 61 49 43 45 62 34 44 56 46 48 38 40 58 37 44
    53 54 59 46 36 42 24 37 47 41 33 62
46 51 75 60 50 41 35 49 41 54 58 49 45 58 41 49 56 63 53 59
    70 30 46 72 62 35 63 43 43 51 37 49 53 58 47 48 45 43 61
    50 53 60 59 78 54 40 44 36 46 78 50
47 39 49 43 53 56 55 52 32 37 60 37 65 53 47 49 55 47 45 53
    65 55 54 56 52 64 54 36 52 63 50 52 34 54 57 47 51 58 60
    54 34 52 51 39 49 60 39 50 42 53 36
48 46 42 50 59 39 50 53 52 31 38 47 45 68 61 47 58 56 53 52
    72 44 45 51 62 44 54 56 54 47 44 30 64 48 57 49 53 63 60
    60 53 51 49 32 56 61 44 45 28 45 46
  
```

DATASIMULATIE II
 ASELEKTE DATA VAN 100 ELEMENTEN BIJ 50 OPERATOREN

VERVOLG

49 44 56 68 59 52 48 45 56 56 46 63 65 45 33 62 44 48 52 53
 45 30 51 40 58 52 57 57 39 32 42 61 64 47 56 41 51 47 54
 58 49 54 52 55 51 60 48 55 59 32 50
 50 65 60 57 41 43 51 54 44 40 72 53 44 57 42 58 78 48 66 42
 45 48 49 55 40 56 46 33 34 51 77 40 52 53 42 48 49 56 54
 41 41 66 56 53 50 59 56 41 57 43 30
 51 54 56 40 50 65 53 53 46 67 46 64 52 54 44 42 44 56 29 44
 50 70 44 49 58 47 44 51 49 41 47 42 77 36 41 39 47 58 38
 69 35 57 69 51 40 43 60 38 53 47 46
 52 50 55 54 57 45 61 60 42 49 58 53 44 30 53 42 67 60 56 50
 45 61 49 49 48 67 46 25 45 54 52 46 76 32 34 49 36 49 52
 57 59 58 47 53 53 62 62 60 47 58 41
 53 55 41 42 43 58 39 35 44 58 54 47 46 68 53 32 52 61 25 55
 55 64 60 42 61 64 49 53 59 62 47 38 48 49 60 49 51 48 55
 52 40 45 32 52 52 57 53 27 34 63 60
 54 48 34 54 69 47 25 53 46 47 49 45 62 64 71 66 57 66 39 35
 53 61 49 42 45 42 64 40 49 52 55 41 55 64 36 54 51 68 55
 55 51 40 48 43 37 43 56 52 47 47 71
 55 44 57 49 35 54 47 54 56 50 59 50 39 51 46 48 45 47 42 59
 54 53 47 57 46 62 57 58 37 51 44 69 53 49 52 50 75 65 42
 52 54 43 55 49 43 46 37 44 26 60 57
 56 66 59 54 69 73 50 56 47 55 50 49 55 76 59 45 43 49 51 41
 42 32 43 54 54 41 41 52 63 48 64 47 61 44 54 41 54 42 64
 50 31 49 69 60 44 73 45 59 55 72 41
 57 47 58 57 28 58 51 23 37 34 61 54 54 53 58 67 44 52 55 50
 55 40 56 55 59 49 56 45 55 43 63 62 58 28 47 74 53 29 39
 40 58 47 53 58 41 63 44 51 35 49 58
 58 32 61 40 45 58 66 42 56 46 75 35 43 36 58 48 50 68 54 44
 46 55 57 47 36 54 50 42 63 40 40 38 48 33 55 55 44 33 54
 54 38 43 60 48 38 47 35 73 33 52 57
 59 46 54 36 51 47 55 48 57 75 51 61 46 35 59 38 62 41 68 55
 49 59 59 50 54 58 45 52 56 68 50 63 65 31 50 55 43 63 54
 54 55 68 63 46 47 44 47 41 37 32 52
 60 59 65 52 43 62 53 39 68 56 54 37 46 55 63 44 38 39 49 50
 39 48 40 44 53 53 48 50 77 58 36 53 41 59 58 49 55 34 41
 57 22 52 44 56 56 54 68 50 64 53 44
 61 43 44 54 38 49 57 45 55 59 64 44 33 50 60 62 62 34 60 41
 44 56 42 46 33 30 45 43 43 34 52 30 53 57 55 48 52 59 59
 58 61 64 52 54 43 27 43 61 42 63 32
 62 48 36 52 53 44 73 54 49 59 54 60 47 48 51 47 46 45 59 38
 55 36 43 52 48 42 61 29 27 53 54 50 28 40 63 59 60 39 42
 52 42 56 36 45 55 39 43 57 53 58 52
 63 50 50 46 45 47 55 55 53 42 45 31 42 71 62 32 60 54 48 48
 53 50 49 50 44 42 40 40 24 34 50 65 48 42 48 40 62 24 38
 36 45 70 53 41 52 40 19 49 31 39 53
 64 35 43 43 55 55 43 37 27 46 59 52 62 59 51 38 55 46 49 49
 64 50 47 48 26 48 48 53 64 47 51 42 59 55 44 55 56 26 57
 55 42 41 43 46 65 46 46 73 36 47 51

DATASIMULATIE II
ASLEKTE DATA VAN 100 ELEMENTEN BIJ 50 OPERATOREN

VERVOLG

```

65 45 50 54 63 42 39 67 66 44 51 49 27 30 40 64 44 42 56 68
    63 51 59 60 41 71 54 50 58 62 45 49 34 59 60 41 53 46 47
    53 50 39 47 51 41 56 33 35 56 49 50
66 57 44 69 34 49 37 41 43 46 61 31 68 36 62 55 50 47 27 35
    44 61 36 62 46 38 51 48 62 65 45 51 44 47 45 46 61 52 38
    53 68 60 52 53 50 64 60 39 43 47 33
67 45 47 45 48 48 53 35 50 47 53 38 58 42 61 63 53 25 55 54
    48 61 43 47 45 54 45 48 33 44 22 56 48 40 49 52 41 63 62
    38 53 47 39 39 58 48 48 60 52 32 52
68 51 51 59 34 51 56 50 58 52 36 57 43 48 56 65 47 65 41 42
    50 40 46 44 60 43 51 54 49 44 56 42 56 47 63 54 45 58 48
    64 36 52 55 52 31 70 43 46 34 31 42
69 49 61 37 44 53 50 35 41 64 40 39 61 30 33 58 70 59 51 52
    57 43 32 63 54 44 30 35 63 47 46 43 40 39 50 45 61 48 54
    50 55 44 64 40 52 61 56 38 35 51 59
70 49 58 41 61 29 68 38 56 49 61 74 48 67 39 48 51 47 52 65
    55 54 56 59 58 63 54 49 53 53 44 43 31 63 47 53 69 36 62
    70 34 46 43 57 44 48 34 50 51 40 44
71 48 27 50 39 44 57 42 41 40 36 47 62 52 44 60 47 64 51 59
    61 54 39 43 52 53 41 51 31 61 34 47 49 63 55 55 60 33 54
    42 60 60 52 65 34 51 53 64 54 39 32
72 62 50 45 35 53 56 55 63 51 55 65 55 55 45 34 44 43 34 32
    56 58 64 52 36 63 56 26 52 51 54 23 46 51 35 68 48 42 37
    43 44 43 70 33 48 47 65 57 49 70 36
73 36 57 56 39 62 51 45 40 63 47 66 31 71 38 51 60 49 53 54
    52 61 40 31 51 44 50 40 44 57 36 49 50 60 62 50 39 34 54
    76 34 65 62 53 49 48 55 46 40 51 55
74 66 35 54 66 42 44 59 55 47 38 58 38 66 52 58 56 55 61 55
    61 52 42 38 33 47 55 40 50 62 30 51 38 61 31 28 57 59 53
    67 58 55 42 46 62 57 47 46 65 47 58
75 58 43 47 49 48 44 64 63 50 38 44 36 54 58 75 46 47 18 49
    53 54 43 66 50 45 62 56 49 38 53 59 47 69 48 57 50 58 40
    69 57 45 55 56 41 63 50 43 54 47 34
76 65 41 54 57 60 28 31 73 38 64 51 45 59 61 53 50 46 65 64
    67 56 54 48 36 34 52 44 59 50 50 42 30 53 59 42 64 58 40
    49 30 48 57 28 56 49 56 45 55 50 43
77 41 70 60 48 46 56 42 53 43 29 30 50 47 50 49 34 34 39 46
    42 53 48 31 43 54 42 46 66 60 58 63 55 44 43 47 64 53 49
    47 39 64 67 39 58 48 61 46 60 55 56
78 58 48 39 52 62 66 66 67 64 38 43 42 50 42 48 32 62 47 54
    50 51 55 52 45 52 59 36 46 39 43 49 62 41 54 41 37 77 71
    47 66 58 55 48 64 57 37 38 51 69 42
79 54 56 52 57 73 31 44 49 55 51 55 37 81 51 59 55 55 49 50
    55 43 44 53 63 51 49 58 66 59 45 54 45 48 50 53 49 31 48
    49 51 81 61 48 38 47 48 67 55 66 31
80 36 51 59 47 55 37 47 40 59 45 36 61 41 59 51 57 50 55 40
    43 54 53 43 51 73 40 51 37 67 40 48 60 40 55 49 39 58 52
    51 45 63 47 34 51 46 60 49 53 47 43

```


DATASIMULATIE II
 AEGELEKTE DATA VAN 100 ELEMENTEN BIJ 50 OPERATOREN

VERVOLG

1 3 31 42 67 45 35 44 39 40 75 43 54 57 50 57 41 59 65 47
 62 58 40 38 48 48 42 53 44 45 53 31 25 67 51 48 30 53 50
 5 49 55 56 51 43 37 36 42 46 34 40
 2 18 45 43 50 37 53 63 49 48 54 56 52 54 37 63 46 49 46 61
 40 48 55 55 44 49 41 54 33 61 53 68 50 65 51 60 51 65 47
 54 50 55 46 45 54 64 55 52 27 38 40
 3 54 59 40 43 53 56 42 44 43 86 47 47 38 40 53 48 56 45 37
 55 46 39 50 40 39 34 44 53 48 40 52 38 38 63 49 27 65 56
 55 40 58 41 47 45 57 42 57 32 44 50
 4 59 42 54 52 58 38 22 49 40 66 66 49 64 44 47 57 61 53 65
 57 43 53 44 51 55 39 50 44 63 67 60 60 40 63 52 50 63 56
 49 71 36 39 32 51 49 63 46 56 48 45
 5 49 59 41 40 58 53 56 47 44 61 45 47 39 60 62 32 57 49 60
 47 55 48 78 57 59 59 54 51 41 55 38 38 51 43 44 71 53 50
 45 62 72 27 48 58 45 53 52 66 43 54
 6 55 35 45 53 51 46 54 48 52 47 36 49 45 38 36 55 47 42 54
 43 53 60 49 27 68 35 55 54 31 49 54 59 43 45 64 52 69 50
 41 55 49 51 36 42 44 60 50 40 45 47
 7 37 64 38 46 52 76 59 43 42 67 45 45 47 60 38 52 47 49 28
 64 65 53 50 44 57 51 49 58 48 52 33 52 57 40 32 45 54 50
 40 57 22 47 59 61 46 51 75 41 39 40
 8 53 43 42 69 30 45 63 49 57 47 44 48 77 61 48 43 54 61 51
 53 34 51 51 30 47 40 35 47 41 47 46 60 58 42 56 55 43 39
 69 32 51 53 53 42 50 40 46 26 59 59
 9 58 50 59 41 51 50 44 48 51 34 38 25 61 53 50 66 51 49 45
 62 51 45 48 30 41 58 48 74 65 45 42 46 59 66 45 72 50 52
 55 41 47 47 53 50 61 65 61 31 45 44
 10 61 47 53 53 51 44 52 63 57 57 51 36 38 55 50 47 37 59 52
 42 46 52 49 70 51 37 50 71 48 51 50 47 46 32 70 31 46 49
 42 51 60 44 49 53 43 42 76 47 41 37
 11 71 34 52 67 53 59 56 31 52 45 36 61 42 50 45 19 48 54 28
 60 45 46 53 51 53 50 60 41 52 41 51 40 33 64 67 41 52 48
 54 68 34 63 60 56 57 41 52 55 41 67
 12 52 43 57 55 52 43 39 44 56 69 42 54 32 35 64 36 26 59 53
 52 49 51 40 49 44 62 39 63 53 54 54 54 43 42 52 32 50 64
 51 50 45 48 32 42 59 45 57 45 55 56
 13 49 37 47 57 55 59 44 72 53 42 53 51 63 51 55 43 66 51 54
 60 35 51 72 56 47 49 48 46 37 56 48 57 46 48 63 49 54 48
 54 53 65 43 66 52 52 43 55 51 53 59
 14 47 66 36 43 63 46 51 70 52 62 51 45 40 47 49 49 50 74 33
 50 50 49 39 51 39 51 57 36 46 39 49 40 56 35 58 61 61 44
 50 60 59 50 42 47 47 41 44 45 56 55
 15 55 61 47 55 39 40 45 43 71 52 55 64 68 49 64 55 47 51 63
 75 49 45 51 63 57 46 53 43 44 42 44 49 66 55 50 56 52 46
 60 64 62 49 45 51 45 30 48 49 57 47
 16 52 63 55 38 65 58 49 47 54 40 35 55 49 50 52 48 53 48 73
 54 59 58 52 51 42 44 36 40 51 63 44 44 45 54 30 47 35 47
 44 45 44 46 52 38 57 43 47 25 52 35

DATASIMULATIE II
 ASELEKTE DATA VAN 100 ELEMENTEN BIJ 50 OPERATOREN

VERVOLG

| | | | | | | | | | | | | | | | | | | | |
|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 87 | 55 | 49 | 36 | 50 | 68 | 53 | 32 | 48 | 46 | 43 | 30 | 71 | 53 | 56 | 40 | 63 | 54 | 54 | 36 |
| | 48 | 57 | 66 | 58 | 35 | 50 | 47 | 63 | 42 | 35 | 61 | 71 | 42 | 65 | 55 | 58 | 46 | 64 | 49 |
| | 48 | 59 | 49 | 43 | 49 | 39 | 38 | 71 | 50 | 54 | 52 | 25 | | | | | | | |
| 95 | 44 | 55 | 52 | 44 | 32 | 53 | 44 | 56 | 61 | 74 | 33 | 40 | 66 | 44 | 37 | 58 | 48 | 50 | 45 |
| | 45 | 26 | 51 | 70 | 57 | 42 | 51 | 55 | 51 | 53 | 43 | 44 | 46 | 33 | 42 | 56 | 35 | 61 | 46 |
| | 49 | 50 | 46 | 53 | 45 | 57 | 42 | 56 | 26 | 23 | 62 | 47 | | | | | | | |
| 99 | 47 | 46 | 50 | 50 | 55 | 54 | 36 | 53 | 48 | 62 | 38 | 65 | 46 | 61 | 48 | 55 | 42 | 33 | 38 |
| | 37 | 50 | 55 | 65 | 54 | 56 | 58 | 54 | 43 | 57 | 45 | 44 | 52 | 64 | 28 | 57 | 50 | 38 | 63 |
| | 55 | 60 | 53 | 49 | 42 | 34 | 43 | 56 | 44 | 59 | 47 | 73 | | | | | | | |
| 100 | 40 | 62 | 52 | 46 | 50 | 43 | 61 | 38 | 46 | 51 | 36 | 48 | 61 | 41 | 42 | 53 | 51 | 47 | 55 |
| | 46 | 54 | 56 | 54 | 48 | 39 | 16 | 30 | 54 | 37 | 39 | 64 | 71 | 62 | 57 | 60 | 47 | 68 | 43 |
| | 45 | 52 | 44 | 64 | 67 | 58 | 56 | 48 | 55 | 59 | 36 | 55 | | | | | | | |

DATASIMULATIE III
 EEN ALGEMENE FAKTOR MET GELIJKE LADINGEN VAN 1,50
 VOOR 30 OPERATOREN BIJ 100 ELEMENTEN

| | | | | | | | | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1 | 60 | 67 | 81 | 84 | 69 | 61 | 85 | 77 | 86 | 65 | 63 | 74 | 64 | 84 | 87 | 72 | 56 | 78 | 70 |
| | 65 | 69 | 81 | 76 | 74 | 53 | 71 | 78 | 59 | 97 | 74 | | | | | | | | |
| | 67 | 63 | 74 | 64 | 65 | 60 | 64 | 62 | 62 | 60 | 86 | 70 | 61 | 67 | 64 | 79 | 59 | 73 | 69 |
| | 90 | 69 | 66 | 75 | 79 | 72 | 74 | 84 | 82 | 82 | 68 | | | | | | | | |
| 2 | 67 | 66 | 72 | 68 | 71 | 78 | 75 | 66 | 69 | 80 | 73 | 64 | 73 | 82 | 75 | 80 | 84 | 64 | 73 |
| | 75 | 62 | 74 | 66 | 69 | 71 | 66 | 49 | 66 | 73 | 71 | | | | | | | | |
| 4 | 75 | 61 | 53 | 90 | 60 | 77 | 69 | 75 | 85 | 57 | 65 | 62 | 68 | 71 | 74 | 67 | 71 | 61 | 68 |
| | 65 | 51 | 53 | 80 | 64 | 79 | 64 | 65 | 76 | 67 | 77 | | | | | | | | |
| 5 | 63 | 58 | 60 | 58 | 90 | 69 | 83 | 88 | 66 | 87 | 68 | 86 | 77 | 67 | 69 | 53 | 74 | 65 | 74 |
| | 77 | 75 | 87 | 69 | 67 | 77 | 75 | 79 | 68 | 55 | 89 | | | | | | | | |
| 6 | 64 | 60 | 70 | 67 | 69 | 74 | 74 | 81 | 69 | 75 | 70 | 87 | 66 | 80 | 63 | 69 | 76 | 79 | 62 |
| | 77 | 91 | 76 | 71 | 96 | 53 | 80 | 88 | 81 | 58 | 97 | | | | | | | | |
| 7 | 53 | 70 | 56 | 77 | 52 | 64 | 72 | 68 | 77 | 60 | 65 | 75 | 66 | 60 | 65 | 61 | 61 | 61 | 74 |
| | 56 | 75 | 53 | 82 | 77 | 74 | 68 | 63 | 60 | 63 | 76 | | | | | | | | |
| 3 | 84 | 64 | 73 | 72 | 68 | 51 | 84 | 66 | 72 | 81 | 68 | 73 | 75 | 75 | 75 | 73 | 69 | 76 | 67 |
| | 74 | 86 | 84 | 93 | 83 | 76 | 77 | 72 | 66 | 72 | 64 | | | | | | | | |
| 9 | 75 | 82 | 68 | 77 | 63 | 72 | 72 | 66 | 64 | 53 | 63 | 88 | 76 | 79 | 67 | 68 | 84 | 74 | 71 |
| | 58 | 79 | 75 | 69 | 61 | 74 | 67 | 59 | 77 | 59 | 83 | | | | | | | | |
| 10 | 56 | 64 | 71 | 70 | 75 | 74 | 67 | 69 | 73 | 83 | 92 | 63 | 68 | 68 | 53 | 53 | 54 | 77 | 71 |
| | 73 | 69 | 64 | 69 | 70 | 74 | 62 | 75 | 87 | 79 | 83 | | | | | | | | |
| 11 | 73 | 71 | 57 | 67 | 50 | 74 | 73 | 57 | 71 | 64 | 82 | 77 | 59 | 74 | 62 | 67 | 64 | 76 | 57 |
| | 79 | 81 | 52 | 55 | 62 | 65 | 61 | 75 | 63 | 64 | 68 | | | | | | | | |
| 12 | 57 | 67 | 81 | 69 | 65 | 72 | 61 | 75 | 60 | 74 | 63 | 63 | 53 | 63 | 57 | 64 | 62 | 65 | 62 |
| | 51 | 47 | 63 | 78 | 64 | 54 | 59 | 82 | 54 | 53 | 44 | | | | | | | | |
| 13 | 63 | 67 | 71 | 81 | 71 | 49 | 65 | 80 | 64 | 48 | 56 | 60 | 63 | 69 | 67 | 59 | 71 | 68 | 69 |
| | 66 | 65 | 58 | 62 | 74 | 74 | 56 | 86 | 65 | 62 | 61 | | | | | | | | |
| 14 | 53 | 67 | 61 | 63 | 53 | 58 | 56 | 73 | 59 | 58 | 69 | 66 | 79 | 60 | 65 | 85 | 65 | 52 | 61 |
| | 55 | 68 | 70 | 57 | 55 | 77 | 62 | 52 | 64 | 73 | 74 | | | | | | | | |
| 15 | 55 | 80 | 61 | 67 | 73 | 72 | 66 | 78 | 72 | 75 | 83 | 71 | 78 | 87 | 72 | 75 | 80 | 51 | 66 |
| | 76 | 79 | 79 | 88 | 81 | 82 | 63 | 67 | 73 | 75 | 81 | | | | | | | | |
| 16 | 80 | 71 | 74 | 80 | 76 | 85 | 75 | 70 | 68 | 90 | 76 | 73 | 78 | 66 | 58 | 78 | 87 | 82 | 82 |
| | 72 | 72 | 74 | 76 | 71 | 76 | 72 | 77 | 73 | 71 | 74 | | | | | | | | |
| 17 | 81 | 63 | 44 | 70 | 46 | 41 | 59 | 67 | 54 | 61 | 65 | 74 | 57 | 65 | 50 | 67 | 75 | 61 | 52 |
| | 67 | 60 | 60 | 71 | 77 | 57 | 54 | 75 | 68 | 74 | 82 | | | | | | | | |
| 18 | 51 | 62 | 62 | 61 | 59 | 64 | 43 | 69 | 69 | 60 | 76 | 50 | 61 | 61 | 72 | 45 | 55 | 52 | 56 |
| | 67 | 76 | 72 | 79 | 71 | 59 | 55 | 61 | 69 | 71 | 64 | | | | | | | | |
| 19 | 61 | 65 | 87 | 85 | 95 | 58 | 75 | 70 | 84 | 61 | 70 | 74 | 89 | 72 | 74 | 67 | 69 | 85 | 77 |
| | 53 | 87 | 79 | 73 | 77 | 79 | 60 | 61 | 76 | 55 | 88 | | | | | | | | |
| 20 | 65 | 70 | 74 | 47 | 64 | 78 | 69 | 65 | 75 | 80 | 66 | 53 | 80 | 66 | 67 | 59 | 79 | 35 | 66 |
| | 62 | 46 | 74 | 75 | 64 | 69 | 77 | 69 | 74 | 63 | 56 | | | | | | | | |
| 21 | 95 | 75 | 86 | 55 | 57 | 66 | 72 | 77 | 82 | 65 | 69 | 54 | 68 | 66 | 76 | 98 | 75 | 74 | 68 |
| | 82 | 73 | 83 | 60 | 77 | 63 | 83 | 74 | 72 | 68 | 53 | | | | | | | | |
| 22 | 70 | 54 | 69 | 65 | 60 | 56 | 62 | 71 | 57 | 67 | 81 | 85 | 85 | 70 | 82 | 80 | 58 | 76 | 59 |
| | 61 | 77 | 77 | 74 | 62 | 77 | 69 | 72 | 52 | 73 | 54 | | | | | | | | |
| 23 | 61 | 69 | 62 | 66 | 77 | 77 | 75 | 84 | 73 | 64 | 57 | 74 | 86 | 93 | 80 | 67 | 75 | 63 | 61 |
| | 79 | 65 | 83 | 82 | 51 | 71 | 78 | 80 | 59 | 66 | 65 | | | | | | | | |
| 24 | 67 | 61 | 66 | 65 | 58 | 76 | 71 | 67 | 59 | 58 | 70 | 58 | 61 | 62 | 50 | 41 | 60 | 62 | 70 |
| | 47 | 59 | 70 | 64 | 68 | 63 | 68 | 60 | 60 | 75 | 62 | | | | | | | | |
| 25 | 68 | 51 | 75 | 59 | 76 | 67 | 78 | 78 | 69 | 73 | 66 | 75 | 73 | 75 | 78 | 67 | 71 | 54 | 80 |
| | 62 | 61 | 78 | 69 | 68 | 85 | 91 | 64 | 82 | 61 | 79 | | | | | | | | |

DATASIMULATIE III
 EEN ALGEMENE FAKTOR MET GELIJKE LADINGEN VAN 0,50
 VOOR 30 OPERATOREN BIJ 100 ELEMENTEN

VERVOLG

| | | | | | | | | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|-----|----|----|----|----|----|----|----|----|----|----|----|
| 25 | 55 | 62 | 67 | 50 | 72 | 67 | 58 | 58 | 53 | 64 | 43 | 60 | 71 | 59 | 64 | 88 | 60 | 63 | 63 |
| | 45 | 70 | 49 | 66 | 69 | 66 | 65 | 58 | 73 | 61 | 76 | | | | | | | | |
| 27 | 47 | 65 | 65 | 68 | 59 | 57 | 62 | 44 | 75 | 56 | 71 | 62 | 67 | 67 | 75 | 71 | 55 | 72 | 79 |
| | 44 | 63 | 60 | 58 | 62 | 68 | 74 | 70 | 72 | 71 | 43 | | | | | | | | |
| 28 | 68 | 75 | 81 | 79 | 78 | 80 | 70 | 66 | 71 | 88 | 68 | 73 | 74 | 82 | 58 | 83 | 72 | 53 | 72 |
| | 74 | 76 | 81 | 71 | 82 | 82 | 69 | 80 | 70 | 65 | 73 | | | | | | | | |
| 29 | 62 | 76 | 69 | 72 | 63 | 58 | 74 | 62 | 57 | 66 | 72 | 83 | 79 | 80 | 59 | 71 | 83 | 71 | 69 |
| | 64 | 72 | 74 | 81 | 68 | 53 | 68 | 60 | 86 | 64 | 71 | | | | | | | | |
| 30 | 56 | 83 | 77 | 62 | 73 | 65 | 74 | 64 | 63 | 83 | 62 | 83 | 66 | 73 | 73 | 61 | 71 | 81 | 65 |
| | 50 | 65 | 69 | 80 | 73 | 67 | 49 | 77 | 68 | 73 | 80 | | | | | | | | |
| 31 | 93 | 79 | 75 | 71 | 79 | 69 | 72 | 100 | 71 | 82 | 82 | 85 | 96 | 94 | 71 | 88 | 70 | 78 | 85 |
| | 70 | 75 | 91 | 90 | 88 | 89 | 54 | 90 | 69 | 76 | 78 | | | | | | | | |
| 32 | 72 | 72 | 75 | 79 | 61 | 81 | 74 | 67 | 63 | 79 | 86 | 90 | 84 | 82 | 62 | 77 | 94 | 79 | 66 |
| | 93 | 73 | 71 | 67 | 77 | 61 | 65 | 80 | 59 | 85 | 73 | | | | | | | | |
| 33 | 63 | 74 | 69 | 57 | 77 | 60 | 62 | 67 | 75 | 79 | 69 | 84 | 81 | 77 | 83 | 70 | 58 | 82 | 78 |
| | 69 | 59 | 78 | 65 | 79 | 81 | 58 | 63 | 59 | 66 | 75 | | | | | | | | |
| 34 | 67 | 79 | 60 | 87 | 62 | 75 | 66 | 79 | 59 | 59 | 80 | 59 | 62 | 79 | 80 | 75 | 77 | 73 | 84 |
| | 89 | 65 | 75 | 89 | 89 | 69 | 76 | 71 | 74 | 74 | 65 | | | | | | | | |
| 35 | 62 | 65 | 74 | 72 | 74 | 65 | 64 | 62 | 80 | 69 | 63 | 62 | 66 | 54 | 61 | 74 | 85 | 69 | 68 |
| | 77 | 72 | 66 | 75 | 80 | 73 | 79 | 60 | 54 | 74 | 59 | | | | | | | | |
| 36 | 72 | 61 | 61 | 66 | 67 | 59 | 69 | 56 | 63 | 58 | 65 | 64 | 73 | 48 | 67 | 79 | 57 | 62 | 70 |
| | 50 | 64 | 46 | 57 | 64 | 55 | 67 | 60 | 72 | 63 | 66 | | | | | | | | |
| 37 | 53 | 53 | 65 | 72 | 64 | 69 | 56 | 77 | 67 | 50 | 56 | 51 | 62 | 53 | 60 | 57 | 61 | 70 | 64 |
| | 71 | 64 | 58 | 62 | 53 | 60 | 79 | 44 | 49 | 57 | 64 | | | | | | | | |
| 38 | 81 | 69 | 67 | 73 | 66 | 78 | 48 | 62 | 77 | 63 | 64 | 84 | 83 | 63 | 82 | 64 | 79 | 59 | 78 |
| | 67 | 76 | 55 | 65 | 70 | 56 | 61 | 67 | 55 | 65 | 74 | | | | | | | | |
| 39 | 68 | 69 | 68 | 69 | 74 | 67 | 67 | 70 | 68 | 62 | 65 | 70 | 95 | 65 | 74 | 78 | 68 | 73 | 71 |
| | 80 | 77 | 78 | 61 | 67 | 72 | 62 | 66 | 78 | 73 | 61 | | | | | | | | |
| 40 | 83 | 65 | 75 | 77 | 81 | 56 | 64 | 85 | 78 | 56 | 78 | 70 | 69 | 70 | 64 | 70 | 73 | 77 | 78 |
| | 71 | 78 | 58 | 66 | 86 | 70 | 62 | 68 | 75 | 72 | 70 | | | | | | | | |
| 41 | 81 | 78 | 76 | 52 | 57 | 69 | 61 | 57 | 64 | 61 | 72 | 81 | 64 | 65 | 82 | 66 | 60 | 65 | 63 |
| | 63 | 61 | 50 | 68 | 57 | 62 | 75 | 78 | 67 | 71 | 70 | | | | | | | | |
| 42 | 70 | 48 | 60 | 82 | 67 | 74 | 71 | 56 | 64 | 75 | 68 | 75 | 71 | 80 | 71 | 52 | 56 | 60 | 60 |
| | 63 | 66 | 72 | 83 | 66 | 79 | 71 | 70 | 65 | 62 | 62 | | | | | | | | |
| 43 | 80 | 68 | 79 | 65 | 82 | 75 | 73 | 80 | 66 | 83 | 65 | 76 | 79 | 68 | 75 | 70 | 75 | 67 | 62 |
| | 80 | 65 | 74 | 73 | 63 | 58 | 68 | 68 | 66 | 73 | 76 | | | | | | | | |
| 44 | 65 | 68 | 48 | 51 | 57 | 55 | 55 | 44 | 57 | 58 | 65 | 61 | 60 | 75 | 50 | 71 | 48 | 66 | 56 |
| | 66 | 53 | 54 | 47 | 51 | 65 | 60 | 59 | 53 | 66 | 57 | | | | | | | | |
| 45 | 64 | 73 | 79 | 63 | 65 | 64 | 67 | 62 | 68 | 78 | 90 | 70 | 85 | 68 | 86 | 82 | 68 | 78 | 68 |
| | 86 | 66 | 76 | 64 | 77 | 82 | 73 | 60 | 70 | 64 | 82 | | | | | | | | |
| 46 | 59 | 62 | 75 | 72 | 65 | 60 | 64 | 74 | 57 | 61 | 61 | 56 | 67 | 65 | 60 | 80 | 48 | 54 | 74 |
| | 53 | 76 | 71 | 81 | 63 | 79 | 64 | 71 | 63 | 67 | 71 | | | | | | | | |
| 47 | 57 | 64 | 65 | 58 | 69 | 43 | 69 | 69 | 70 | 56 | 56 | 55 | 45 | 76 | 61 | 63 | 59 | 64 | 43 |
| | 62 | 53 | 61 | 65 | 57 | 42 | 47 | 58 | 52 | 56 | 49 | | | | | | | | |
| 48 | 69 | 78 | 66 | 70 | 78 | 82 | 78 | 64 | 63 | 71 | 75 | 45 | 69 | 58 | 61 | 58 | 61 | 61 | 81 |
| | 86 | 61 | 70 | 71 | 72 | 66 | 60 | 77 | 78 | 58 | 56 | | | | | | | | |
| 49 | 89 | 72 | 61 | 70 | 67 | 66 | 57 | 40 | 53 | 75 | 65 | 66 | 79 | 74 | 56 | 63 | 59 | 72 | 69 |
| | 78 | 66 | 71 | 59 | 61 | 61 | 71 | 66 | 63 | 74 | 69 | | | | | | | | |
| 50 | 68 | 73 | 71 | 70 | 70 | 81 | 76 | 70 | 61 | 61 | 71 | 75 | 74 | 72 | 65 | 75 | 67 | 69 | 80 |
| | 74 | 82 | 78 | 75 | 69 | 80 | 71 | 65 | 66 | 70 | 66 | | | | | | | | |

DATASIMULATIE III
 EEN ALGEMENE FAKTOR MET GELIJKE LADINGEN VAN 0,50
 VOOR 30 OPERATOREN BIJ 100 ELEMENTEN

VERVULG

| | | | | | | | | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 51 | 53 | 63 | 69 | 58 | 60 | 48 | 52 | 68 | 57 | 70 | 65 | 56 | 68 | 45 | 63 | 53 | 47 | 54 | 74 |
| | 59 | 72 | 71 | 48 | 78 | 65 | 54 | 54 | 53 | 64 | 46 | | | | | | | | |
| 52 | 57 | 72 | 61 | 76 | 57 | 61 | 72 | 81 | 81 | 72 | 77 | 64 | 57 | 59 | 78 | 79 | 82 | 75 | 67 |
| | 71 | 75 | 62 | 76 | 69 | 78 | 83 | 79 | 82 | 65 | 66 | | | | | | | | |
| 53 | 71 | 58 | 60 | 71 | 58 | 60 | 73 | 66 | 67 | 62 | 76 | 70 | 62 | 66 | 67 | 83 | 79 | 59 | 74 |
| | 55 | 69 | 76 | 77 | 77 | 78 | 82 | 71 | 55 | 66 | 90 | | | | | | | | |
| 54 | 65 | 79 | 80 | 72 | 68 | 70 | 65 | 61 | 68 | 55 | 64 | 74 | 76 | 77 | 53 | 64 | 69 | 62 | 75 |
| | 66 | 61 | 61 | 49 | 58 | 47 | 63 | 74 | 62 | 67 | 68 | | | | | | | | |
| 55 | 74 | 73 | 85 | 73 | 81 | 77 | 77 | 81 | 79 | 64 | 81 | 82 | 69 | 84 | 94 | 78 | 95 | 68 | 79 |
| | 66 | 53 | 74 | 73 | 69 | 79 | 80 | 66 | 77 | 60 | 75 | | | | | | | | |
| 56 | 67 | 79 | 60 | 73 | 54 | 62 | 68 | 57 | 55 | 64 | 62 | 60 | 75 | 79 | 54 | 69 | 52 | 84 | 78 |
| | 64 | 64 | 54 | 59 | 55 | 60 | 71 | 49 | 67 | 60 | 85 | | | | | | | | |
| 57 | 59 | 73 | 59 | 50 | 63 | 62 | 51 | 59 | 54 | 61 | 57 | 75 | 77 | 61 | 49 | 52 | 63 | 59 | 60 |
| | 73 | 44 | 56 | 47 | 67 | 52 | 83 | 53 | 65 | 67 | 67 | | | | | | | | |
| 58 | 68 | 66 | 70 | 74 | 69 | 69 | 75 | 73 | 82 | 74 | 75 | 73 | 68 | 71 | 84 | 53 | 53 | 71 | 72 |
| | 65 | 76 | 73 | 80 | 80 | 53 | 54 | 64 | 56 | 77 | 68 | | | | | | | | |
| 59 | 81 | 69 | 78 | 80 | 72 | 72 | 78 | 90 | 78 | 81 | 69 | 91 | 68 | 73 | 65 | 69 | 71 | 75 | 63 |
| | 62 | 69 | 57 | 80 | 73 | 64 | 73 | 74 | 67 | 57 | 68 | | | | | | | | |
| 60 | 65 | 55 | 62 | 55 | 83 | 48 | 68 | 63 | 65 | 46 | 64 | 55 | 66 | 52 | 62 | 65 | 69 | 70 | 58 |
| | 62 | 64 | 70 | 65 | 57 | 75 | 51 | 73 | 64 | 76 | 48 | | | | | | | | |
| 61 | 57 | 75 | 70 | 70 | 58 | 54 | 58 | 50 | 68 | 66 | 52 | 54 | 62 | 66 | 76 | 63 | 60 | 67 | 65 |
| | 70 | 58 | 61 | 69 | 55 | 56 | 77 | 64 | 66 | 77 | 66 | | | | | | | | |
| 62 | 69 | 65 | 46 | 57 | 55 | 62 | 78 | 50 | 66 | 79 | 75 | 49 | 75 | 63 | 53 | 57 | 65 | 66 | 55 |
| | 65 | 75 | 45 | 62 | 71 | 70 | 45 | 60 | 72 | 65 | 65 | | | | | | | | |
| 63 | 77 | 74 | 74 | 73 | 61 | 58 | 77 | 74 | 80 | 80 | 56 | 85 | 64 | 55 | 61 | 73 | 66 | 71 | 79 |
| | 70 | 74 | 69 | 74 | 69 | 73 | 69 | 69 | 77 | 73 | 65 | | | | | | | | |
| 64 | 64 | 74 | 72 | 89 | 68 | 72 | 73 | 71 | 75 | 63 | 62 | 70 | 69 | 65 | 75 | 57 | 67 | 74 | 75 |
| | 79 | 71 | 73 | 70 | 67 | 76 | 67 | 61 | 69 | 78 | 64 | | | | | | | | |
| 65 | 61 | 46 | 51 | 62 | 62 | 75 | 67 | 53 | 38 | 54 | 57 | 70 | 57 | 46 | 59 | 69 | 55 | 69 | 46 |
| | 81 | 52 | 46 | 65 | 64 | 61 | 57 | 55 | 52 | 68 | 52 | | | | | | | | |
| 66 | 81 | 73 | 54 | 70 | 58 | 66 | 73 | 63 | 61 | 58 | 75 | 73 | 65 | 79 | 63 | 66 | 69 | 68 | 72 |
| | 56 | 77 | 76 | 65 | 73 | 61 | 68 | 52 | 65 | 77 | 69 | | | | | | | | |
| 67 | 56 | 59 | 73 | 61 | 56 | 69 | 66 | 66 | 68 | 65 | 81 | 67 | 52 | 65 | 65 | 64 | 77 | 68 | 62 |
| | 72 | 54 | 60 | 72 | 65 | 51 | 73 | 55 | 48 | 60 | 69 | | | | | | | | |
| 68 | 62 | 67 | 79 | 54 | 65 | 61 | 67 | 66 | 73 | 51 | 62 | 67 | 66 | 60 | 62 | 55 | 62 | 60 | 81 |
| | 58 | 69 | 76 | 58 | 67 | 55 | 62 | 52 | 59 | 55 | 71 | | | | | | | | |
| 69 | 64 | 66 | 63 | 59 | 73 | 66 | 63 | 62 | 76 | 86 | 69 | 89 | 70 | 82 | 73 | 71 | 78 | 83 | 82 |
| | 72 | 81 | 69 | 72 | 61 | 67 | 78 | 86 | 75 | 73 | 65 | | | | | | | | |
| 70 | 73 | 31 | 76 | 59 | 63 | 69 | 83 | 54 | 76 | 71 | 55 | 75 | 78 | 81 | 48 | 66 | 67 | 72 | 71 |
| | 79 | 88 | 65 | 75 | 56 | 79 | 62 | 53 | 64 | 68 | 85 | | | | | | | | |
| 71 | 65 | 73 | 64 | 67 | 70 | 64 | 61 | 74 | 65 | 77 | 52 | 74 | 59 | 52 | 60 | 56 | 68 | 47 | 69 |
| | 82 | 82 | 63 | 57 | 67 | 53 | 65 | 64 | 71 | 44 | 54 | | | | | | | | |
| 72 | 79 | 59 | 55 | 74 | 72 | 55 | 64 | 68 | 67 | 71 | 84 | 62 | 71 | 61 | 69 | 61 | 83 | 62 | 68 |
| | 76 | 83 | 58 | 67 | 69 | 75 | 69 | 63 | 77 | 81 | 53 | | | | | | | | |
| 73 | 68 | 68 | 82 | 79 | 65 | 64 | 62 | 73 | 61 | 69 | 73 | 54 | 69 | 70 | 62 | 75 | 61 | 76 | 69 |
| | 68 | 65 | 67 | 71 | 78 | 70 | 71 | 72 | 80 | 66 | 69 | | | | | | | | |
| 74 | 65 | 66 | 63 | 76 | 71 | 60 | 78 | 66 | 65 | 80 | 71 | 82 | 31 | 60 | 78 | 86 | 55 | 70 | 54 |
| | 66 | 75 | 70 | 73 | 59 | 59 | 51 | 56 | 64 | 79 | 69 | | | | | | | | |
| 75 | 70 | 61 | 75 | 57 | 60 | 69 | 54 | 52 | 59 | 43 | 69 | 48 | 68 | 51 | 65 | 87 | 70 | 55 | 75 |
| | 76 | 49 | 73 | 55 | 57 | 64 | 75 | 73 | 52 | 51 | 64 | | | | | | | | |

DATASIMULATI. III
 EEN ALGEMENE FAKTOR MET GELIJKE LADINGEN VAN 1,50
 VOOR 40 OPERATOREN BIJ 100 ELEMENTEN

VOLG

| | | | | | | | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|-----|----|----|----|----|----|----|----|
| 10 | 1 | 72 | 75 | 54 | 55 | 63 | 65 | 64 | 65 | 65 | 68 | 64 | 74 | 60 | 56 | 59 | 77 | 72 |
| | 56 | 71 | 44 | 61 | 70 | 61 | 58 | 55 | 56 | 72 | 57 | | | | | | | |
| 11 | 55 | 64 | 53 | 64 | 55 | 43 | 54 | 50 | 73 | 61 | 72 | 67 | 64 | 62 | 51 | 72 | 63 | 54 |
| | 72 | 52 | 56 | 60 | 61 | 51 | 60 | 52 | 63 | 48 | 63 | | | | | | | |
| 12 | 71 | 64 | 53 | 65 | 64 | 61 | 74 | 63 | 48 | 64 | 53 | 68 | 65 | 65 | 61 | 76 | 71 | 65 |
| | 63 | 70 | 62 | 74 | 64 | 75 | 80 | 55 | 65 | 60 | 53 | | | | | | | |
| 13 | 77 | 80 | 69 | 71 | 73 | 78 | 71 | 71 | 65 | 82 | 68 | 73 | 77 | 73 | 71 | 75 | 57 | 73 |
| | 62 | 69 | 81 | 75 | 71 | 77 | 65 | 79 | 66 | 67 | 67 | | | | | | | |
| 14 | 65 | 66 | 67 | 56 | 72 | 74 | 72 | 61 | 55 | 61 | 80 | 63 | 81 | 66 | 75 | 62 | 64 | 72 |
| | 75 | 75 | 77 | 73 | 59 | 63 | 66 | 64 | 48 | 49 | 64 | | | | | | | |
| 15 | 62 | 65 | 87 | 73 | 86 | 68 | 61 | 84 | 75 | 81 | 103 | 75 | 70 | 62 | 89 | 81 | 76 | 75 |
| | 65 | 72 | 67 | 62 | 91 | 86 | 80 | 60 | 75 | 77 | 80 | | | | | | | |
| 16 | 63 | 66 | 65 | 66 | 66 | 78 | 73 | 66 | 70 | 64 | 72 | 76 | 75 | 68 | 70 | 79 | 81 | 67 |
| | 60 | 72 | 61 | 55 | 69 | 53 | 73 | 92 | 83 | 67 | 61 | | | | | | | |
| 17 | 57 | 69 | 77 | 70 | 65 | 67 | 53 | 63 | 52 | 70 | 71 | 65 | 77 | 78 | 55 | 63 | 53 | 65 |
| | 71 | 75 | 74 | 57 | 55 | 62 | 65 | 78 | 64 | 68 | 65 | | | | | | | |
| 18 | 72 | 83 | 54 | 78 | 69 | 83 | 67 | 65 | 79 | 76 | 69 | 77 | 83 | 64 | 85 | 90 | 72 | 92 |
| | 73 | 56 | 74 | 69 | 62 | 87 | 59 | 73 | 76 | 82 | 60 | | | | | | | |
| 19 | 61 | 61 | 72 | 77 | 65 | 73 | 67 | 78 | 71 | 73 | 65 | 72 | 64 | 64 | 64 | 69 | 79 | 78 |
| | 65 | 77 | 58 | 60 | 81 | 75 | 60 | 59 | 77 | 71 | 70 | | | | | | | |
| 20 | 65 | 75 | 89 | 57 | 57 | 68 | 74 | 76 | 74 | 63 | 50 | 52 | 62 | 65 | 77 | 62 | 55 | 61 |
| | 75 | 70 | 74 | 47 | 63 | 56 | 65 | 72 | 65 | 56 | 53 | | | | | | | |
| 21 | 62 | 60 | 55 | 61 | 61 | 73 | 55 | 65 | 60 | 68 | 59 | 65 | 58 | 67 | 52 | 54 | 50 | 64 |
| | 65 | 52 | 72 | 66 | 63 | 79 | 77 | 78 | 71 | 61 | 59 | | | | | | | |
| 22 | 53 | 66 | 61 | 58 | 74 | 73 | 70 | 48 | 64 | 70 | 65 | 59 | 64 | 59 | 53 | 66 | 75 | 82 |
| | 57 | 68 | 59 | 58 | 60 | 62 | 48 | 68 | 72 | 57 | 62 | | | | | | | |
| 23 | 83 | 54 | 57 | 51 | 71 | 66 | 52 | 73 | 54 | 76 | 75 | 70 | 56 | 77 | 70 | 66 | 68 | 74 |
| | 55 | 67 | 50 | 71 | 73 | 73 | 73 | 68 | 79 | 74 | 67 | | | | | | | |
| 24 | 65 | 60 | 76 | 64 | 67 | 67 | 75 | 66 | 67 | 70 | 52 | 70 | 62 | 75 | 65 | 52 | 68 | 51 |
| | 72 | 66 | 75 | 66 | 59 | 74 | 72 | 70 | 58 | 55 | 68 | | | | | | | |
| 25 | 66 | 53 | 63 | 64 | 70 | 84 | 69 | 79 | 80 | 56 | 65 | 70 | 65 | 63 | 74 | 83 | 73 | 59 |
| | 67 | 69 | 76 | 62 | 76 | 75 | 75 | 81 | 73 | 66 | | | | | | | | |
| 26 | 51 | 70 | 75 | 57 | 67 | 67 | 77 | 67 | 66 | 63 | 70 | 67 | 67 | 51 | 73 | 66 | 64 | 54 |
| | 68 | 57 | 75 | 71 | 57 | 79 | 79 | 50 | 64 | 61 | 52 | | | | | | | |
| 27 | 79 | 66 | 70 | 72 | 68 | 69 | 74 | 83 | 68 | 76 | 72 | 77 | 70 | 59 | 59 | 60 | 64 | 68 |
| | 53 | 77 | 83 | 73 | 61 | 82 | 77 | 71 | 72 | 83 | 76 | | | | | | | |
| 28 | 59 | 65 | 61 | 42 | 60 | 73 | 75 | 50 | 61 | 71 | 62 | 70 | 65 | 71 | 64 | 62 | 64 | 67 |
| | 62 | 69 | 60 | 63 | 74 | 65 | 61 | 69 | 63 | 65 | 43 | | | | | | | |
| 29 | 52 | 66 | 57 | 62 | 86 | 68 | 65 | 68 | 53 | 80 | 62 | 60 | 63 | 78 | 64 | 66 | 65 | 70 |
| | 66 | 64 | 52 | 83 | 70 | 72 | 51 | 70 | 69 | 70 | 79 | | | | | | | |
| 30 | 56 | 75 | 72 | 62 | 68 | 59 | 73 | 55 | 76 | 70 | 72 | 61 | 69 | 56 | 60 | 72 | 75 | 69 |
| | 65 | 53 | 54 | 74 | 71 | 65 | 55 | 75 | 72 | 70 | 74 | | | | | | | |
| 31 | 57 | 62 | 56 | 70 | 51 | 60 | 65 | 63 | 71 | 69 | 63 | 61 | 66 | 61 | 51 | 60 | 58 | 57 |
| | 67 | 65 | 56 | 67 | 63 | 68 | 71 | 73 | 51 | 68 | 58 | | | | | | | |
| 32 | 72 | 67 | 60 | 61 | 65 | 63 | 56 | 74 | 63 | 78 | 60 | 64 | 67 | 72 | 51 | 50 | 75 | 55 |
| | 49 | 65 | 67 | 60 | 72 | 62 | 56 | 76 | 59 | 70 | 66 | | | | | | | |
| 33 | 64 | 61 | 60 | 58 | 68 | 63 | 53 | 51 | 66 | 57 | 73 | 75 | 45 | 62 | 74 | 82 | 57 | 50 |
| | 57 | 46 | 60 | 53 | 55 | 75 | 63 | 60 | 38 | 53 | 73 | | | | | | | |
| 34 | 56 | 72 | 65 | 60 | 65 | 70 | 59 | 47 | 66 | 65 | 58 | 81 | 65 | 74 | 63 | 66 | 53 | 75 |
| | 65 | 68 | 60 | 61 | 75 | 56 | 78 | 42 | 58 | 71 | 53 | | | | | | | |

DATASIMULATIE IV
EEN ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
(ZIE SIMULATIE IV POPULATIE FAKTORSTRUCTUUR)
VOOR 27 OPERATOREN BIJ 600 ELEMENTEN

VF/VOL ,

[illegible]

DATASIMULATIE IV
EEN ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
(ZIE SIMULATIE IV POPULATIE FAKTORSTRUCTUUR)
VOOR 27 OPERATOREN BIJ 400 ELEMENTEN

VERVOLG

| | | | | | | | | | | | | | | | | | | | |
|----|-----|-----|-----|-----|------|-----|-----|----|-----|----|----|----|----|-----|-----|----|-----|----|-----|
| 77 | 72 | 74 | 81 | 69 | 14 | 72 | 76 | 22 | 75 | 82 | 74 | 58 | 53 | 74 | 81 | 76 | 72 | 55 | 92 |
| | 57 | 45 | 55 | 33 | 55 | 71 | 67 | 70 | | | | | | | | | | | |
| 77 | 97 | 87 | 91 | 99 | 31 | 85 | 84 | 19 | 99 | 98 | 85 | 68 | 82 | 100 | 97 | 64 | 86 | 65 | 85 |
| | 83 | 51 | 65 | 31 | 74 | 78 | 85 | 62 | | | | | | | | | | | |
| 78 | 85 | 99 | 101 | 96 | 35 | 86 | 97 | 30 | 95 | 95 | 94 | 72 | 87 | 95 | 106 | 65 | 90 | 84 | 96 |
| | 89 | 57 | 80 | 51 | 80 | 79 | 85 | 81 | | | | | | | | | | | |
| 79 | 102 | 101 | 99 | 79 | 26 | 103 | 87 | 28 | 103 | 83 | 39 | 79 | 91 | 85 | 89 | 78 | 104 | 97 | 89 |
| | 65 | 48 | 75 | 49 | 70 | 79 | 99 | 68 | | | | | | | | | | | |
| 80 | 76 | 74 | 80 | 74 | 19 | 86 | 70 | 26 | 92 | 92 | 80 | 68 | 59 | 61 | 82 | 72 | 85 | 65 | 67 |
| | 73 | 35 | 66 | 54 | 52 | 58 | 74 | 69 | | | | | | | | | | | |
| 81 | 91 | 90 | 82 | 93 | 43 | 86 | 84 | 27 | 85 | 83 | 92 | 72 | 74 | 87 | 93 | 79 | 34 | 39 | 86 |
| | 83 | 41 | 66 | 38 | 64 | 78 | 81 | 70 | | | | | | | | | | | |
| 82 | 67 | 80 | 67 | 102 | 12 | 83 | 81 | 15 | 55 | 86 | 90 | 77 | 65 | 86 | 91 | 69 | 62 | 82 | 85 |
| | 87 | 59 | 66 | 39 | 65 | 65 | 66 | 63 | | | | | | | | | | | |
| 82 | 88 | 97 | 81 | 91 | 32 | 75 | 89 | 32 | 94 | 94 | 90 | 74 | 75 | 93 | 108 | 68 | 39 | 85 | 85 |
| | 95 | 43 | 85 | 52 | 74 | 76 | 86 | 67 | | | | | | | | | | | |
| 83 | 83 | 75 | 89 | 98 | 52 | 93 | 105 | 34 | 104 | 87 | 84 | 78 | 81 | 109 | 92 | 76 | 103 | 71 | 109 |
| | 66 | 51 | 86 | 58 | 87 | 84 | 91 | 82 | | | | | | | | | | | |
| 85 | 81 | 101 | 85 | 72 | 16 | 82 | 77 | 27 | 68 | 81 | 89 | 70 | 71 | 86 | 85 | 73 | 81 | 80 | 82 |
| | 97 | 47 | 74 | 42 | 69 | 74 | 77 | 75 | | | | | | | | | | | |
| 86 | 84 | 63 | 77 | 65 | 25 | 93 | 96 | 25 | 102 | 91 | 82 | 68 | 61 | 75 | 97 | 52 | 80 | 70 | 67 |
| | 99 | 63 | 59 | 41 | 65 | 59 | 85 | 67 | | | | | | | | | | | |
| 87 | 64 | 70 | 79 | 75 | 24 | 72 | 85 | 30 | 89 | 96 | 79 | 53 | 69 | 70 | 83 | 70 | 83 | 51 | 67 |
| | 79 | 33 | 45 | 22 | 62 | 61 | 76 | 62 | | | | | | | | | | | |
| 88 | 72 | 76 | 72 | 83 | 33 | 82 | 86 | 26 | 85 | 83 | 82 | 64 | 54 | 77 | 81 | 66 | 75 | 87 | 69 |
| | 78 | 33 | 55 | 39 | 68 | 60 | 72 | 63 | | | | | | | | | | | |
| 89 | 83 | 94 | 85 | 85 | 28 | 89 | 96 | 45 | 91 | 99 | 82 | 69 | 79 | 84 | 89 | 65 | 90 | 90 | 82 |
| | 95 | 48 | 71 | 34 | 69 | 72 | 80 | 68 | | | | | | | | | | | |
| 90 | 71 | 80 | 84 | 68 | 25 | 72 | 74 | 24 | 72 | 58 | 80 | 67 | 75 | 65 | 74 | 70 | 62 | 60 | 75 |
| | 85 | 41 | 71 | 43 | 75 | 58 | 76 | 81 | | | | | | | | | | | |
| 91 | 99 | 96 | 83 | 103 | 38 | 86 | 83 | 37 | 93 | 85 | 81 | 90 | 82 | 93 | 93 | 76 | 98 | 81 | 95 |
| | 102 | 63 | 79 | 70 | 84 | 83 | 95 | 94 | | | | | | | | | | | |
| 92 | 89 | 95 | 102 | 77 | 32</ | | | | | | | | | | | | | | |

1. ATASIMULATIE IV
 2. ALG. ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
 3. 2E SIMULATIE IV POPULATIE FAKTORSTRUKTUUR
 VOOR 27 OPERATOREN BIJ 600 ELEMENTEN

V L K V J L G

[illegible]

DATA SIMULATIE IV
 EEN ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
 (21e SIMULATIE IV POPULATIE FAKTORSTRUKTUUR)
 VOOR 27 OPERATIEËN AUF 600 ELEMENTEN

VERVOLG

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|----|-----|----|----|----|-----|----|----|-----|-----|----|----|-----|-----|----|-----|-----|-----|
| 156 | 53 | 59 | 69 | 35 | 70 | 53 | 87 | 79 | 87 | 75 | 74 | 72 | 6 | 51 | 19 | 69 | 57 | 91 | 78 |
| 157 | 53 | 36 | 8 | 15 | 17 | 72 | 100 | 20 | 88 | 94 | 83 | 70 | 70 | 12 | 85 | 73 | 84 | 70 | 80 |
| 158 | 73 | 12 | 60 | 43 | 55 | 57 | 91 | 60 | 80 | 70 | 58 | 61 | 41 | 72 | 70 | 52 | 73 | 46 | 12 |
| 159 | 53 | 75 | 70 | 66 | 20 | 69 | 50 | 33 | 80 | 70 | 58 | 61 | 41 | 72 | 70 | 52 | 73 | 46 | 12 |
| 160 | 50 | 35 | 56 | 32 | 53 | 66 | 74 | 55 | 81 | 85 | 90 | 57 | 30 | 58 | 84 | 22 | 77 | 94 | 88 |
| 161 | 51 | 46 | 70 | 43 | 58 | 79 | 84 | 78 | 51 | 46 | 70 | 43 | 58 | 79 | 84 | 78 | 51 | 46 | 70 |
| 162 | 70 | 55 | 24 | 70 | 25 | 83 | 70 | 24 | 72 | 75 | 92 | 64 | 65 | 11 | 74 | 60 | 31 | 54 | 74 |
| 163 | 84 | 18 | 61 | 43 | 66 | 68 | 67 | 61 | 70 | 51 | 78 | 89 | 18 | 24 | 78 | 32 | 87 | 87 | 87 |
| 164 | 60 | 34 | 70 | 31 | 65 | 65 | 86 | 81 | 60 | 34 | 70 | 31 | 65 | 65 | 86 | 81 | 60 | 34 | 70 |
| 165 | 83 | 89 | 37 | 18 | 29 | 37 | 87 | 23 | 93 | 67 | 85 | 67 | 62 | 72 | 92 | 65 | 40 | 48 | 76 |
| 166 | 67 | 36 | 60 | 42 | 70 | 68 | 83 | 31 | 62 | 96 | 70 | 17 | 21 | 79 | 89 | 33 | 72 | 92 | 82 |
| 167 | 75 | 30 | 70 | 37 | 70 | 66 | 67 | 62 | 81 | 35 | 85 | 85 | 50 | 65 | 63 | 85 | 75 | 30 | 70 |
| 168 | 50 | 35 | 92 | 80 | 49 | 98 | 91 | 29 | 59 | 35 | 98 | 73 | 75 | 16 | 61 | 85 | 100 | 70 | 100 |
| 169 | 92 | 43 | 84 | 59 | 79 | 74 | 101 | 80 | 37 | 99 | 70 | 81 | 26 | 86 | 100 | 20 | 31 | 101 | 79 |
| 170 | 85 | 47 | 80 | 36 | 65 | 65 | 86 | 63 | 74 | 84 | 72 | 97 | 70 | 97 | 67 | 85 | 74 | 88 | 88 |
| 171 | 74 | 33 | 77 | 88 | 10 | 56 | 88 | 23 | 3 | 70 | 76 | 75 | 65 | 91 | 85 | 80 | 10 | 67 | 95 |
| 172 | 97 | 33 | 61 | 55 | 64 | 76 | 89 | 74 | 72 | 92 | 82 | 81 | 35 | 85 | 85 | 50 | 65 | 63 | 85 |
| 173 | 72 | 92 | 81 | 81 | 8 | 32 | 81 | 23 | 80 | 89 | 90 | 61 | 65 | 73 | 95 | 65 | 63 | 72 | 88 |
| 174 | 89 | 36 | 74 | 46 | 64 | 68 | 77 | 70 | 97 | 95 | 98 | 93 | 38 | 71 | 74 | 70 | 65 | 63 | 76 |
| 175 | 95 | 27 | 55 | 40 | 62 | 70 | 77 | 47 | 95 | 27 | 55 | 40 | 62 | 70 | 77 | 47 | 95 | 27 | 55 |
| 176 | 73 | 92 | 90 | 82 | 15 | 55 | 90 | 11 | 62 | 104 | 86 | 33 | 79 | 106 | 87 | 68 | 82 | 66 | 94 |
| 177 | 114 | 37 | 63 | 45 | 75 | 84 | 82 | 74 | 93 | 91 | 80 | 86 | 15 | 79 | 80 | 31 | 84 | 95 | 96 |
| 178 | 73 | 40 | 68 | 42 | 67 | 62 | 87 | 57 | 61 | 65 | 96 | 81 | 60 | 81 | 87 | 74 | 93 | 69 | 90 |
| 179 | 82 | 90 | 60 | 91 | 22 | 93 | 77 | 15 | 72 | 92 | 84 | 56 | 76 | 92 | 91 | 78 | 91 | 65 | 75 |
| 180 | 83 | 46 | 60 | 57 | 71 | 64 | 76 | 67 | 71 | 85 | 50 | 56 | 59 | 79 | 80 | 56 | 72 | 53 | 65 |
| 181 | 60 | 26 | 65 | 18 | 61 | 57 | 55 | 55 | 65 | 74 | 77 | 88 | 32 | 64 | 81 | 14 | 80 | 83 | 80 |
| 182 | 72 | 51 | 60 | 34 | 72 | 57 | 62 | 62 | 80 | 83 | 80 | 66 | 47 | 73 | 84 | 65 | 65 | 64 | 79 |
| 183 | 92 | 84 | 90 | 84 | 25 | 90 | 96 | 36 | 89 | 103 | 89 | 72 | 80 | 79 | 91 | 75 | 85 | 72 | 71 |
| 184 | 92 | 57 | 51 | 40 | 74 | 62 | 79 | 67 | 83 | 90 | 86 | 56 | 61 | 92 | 105 | 88 | 53 | 74 | 97 |
| 185 | 83 | 90 | 61 | 90 | 22 | 80 | 88 | 38 | 73 | 90 | 86 | 56 | 61 | 92 | 105 | 88 | 53 | 74 | 97 |
| 186 | 89 | 57 | 84 | 56 | 78 | 82 | 83 | 78 | 77 | 79 | 101 | 61 | 79 | 105 | 78 | 89 | 53 | 81 | 96 |
| 187 | 69 | 55 | 35 | 58 | 75 | 80 | 83 | 84 | 82 | 80 | 63 | 74 | 60 | 81 | 84 | 72 | 85 | 67 | 72 |
| 188 | 82 | 71 | 66 | 77 | 23 | 95 | 82 | 34 | 89 | 80 | 63 | 74 | 60 | 81 | 84 | 72 | 85 | 67 | 72 |
| 189 | 78 | 54 | 85 | 41 | 47 | 61 | 86 | 70 | 76 | 73 | 66 | 83 | 36 | 86 | 91 | 44 | 77 | 88 | 82 |
| 190 | 70 | 45 | 59 | 45 | 59 | 76 | 77 | 81 | 69 | 63 | 61 | 90 | 86 | 79 | 68 | 86 | 70 | 45 | 59 |
| 191 | 75 | 80 | 105 | 80 | 21 | 91 | 64 | 20 | 66 | 82 | 87 | 70 | 60 | 63 | 66 | 68 | 61 | 72 | 73 |
| 192 | 87 | 26 | 64 | 50 | 70 | 60 | 72 | 72 | 89 | 68 | 92 | 79 | 21 | 85 | 76 | 19 | 71 | 73 | 56 |
| 193 | 63 | 16 | 63 | 23 | 57 | 59 | 67 | 50 | 79 | 71 | 73 | 56 | 58 | 87 | 84 | 63 | 71 | 59 | 74 |

DATA SIMULATIE IV
EEN ALGEMENE FACTOR EN VIJF SUBFACTOREN
(ZIE SIMULATIE IV POPULATIE FACTORSTRUCTUUR)
VOOR 27 OPERATIEËN BIJ 60° ELEMENTEN

VEF VOLG

[illegible]

DATASIMULATIE IV
EEN ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
(ZIE SIMULATIE IV POPULATIE FAKTORSTRUKTUUR)
VOOR 27 OPERATOREN BIJ 600 ELEMENTEN

VERVOLG

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|----|----|-----|----|-----|-----|----|-----|----|-----|----|----|----|-----|----|-----|----|-----|
| 251 | 87 | 75 | 85 | 85 | 40 | 112 | 84 | 36 | 91 | 85 | 88 | 99 | 74 | 87 | 83 | 83 | 99 | 77 | 92 |
| | 79 | 51 | 88 | 63 | 80 | 83 | 90 | 78 | | | | | | | | | | | |
| 252 | 85 | 85 | 95 | 73 | 28 | 82 | 68 | 21 | 87 | 85 | 93 | 67 | 79 | 77 | 86 | 75 | 82 | 82 | 75 |
| | 86 | 33 | 69 | 49 | 84 | 72 | 85 | 72 | | | | | | | | | | | |
| 253 | 72 | 93 | 94 | 91 | 20 | 108 | 103 | 24 | 78 | 95 | 81 | 68 | 73 | 82 | 103 | 64 | 74 | 69 | 92 |
| | 90 | 27 | 72 | 39 | 53 | 73 | 69 | 62 | | | | | | | | | | | |
| 254 | 58 | 80 | 89 | 73 | 21 | 89 | 87 | 31 | 78 | 74 | 87 | 65 | 68 | 84 | 86 | 78 | 68 | 75 | 85 |
| | 84 | 30 | 61 | 35 | 59 | 61 | 66 | 51 | | | | | | | | | | | |
| 255 | 80 | 78 | 92 | 86 | 12 | 84 | 64 | 23 | 60 | 78 | 97 | 71 | 70 | 88 | 73 | 46 | 80 | 69 | 94 |
| | 82 | 30 | 59 | 43 | 68 | 75 | 72 | 65 | | | | | | | | | | | |
| 256 | 91 | 73 | 90 | 76 | 23 | 82 | 74 | 21 | 96 | 75 | 94 | 73 | 59 | 76 | 74 | 62 | 100 | 65 | 77 |
| | 66 | 27 | 65 | 49 | 61 | 66 | 89 | 67 | | | | | | | | | | | |
| 257 | 85 | 83 | 89 | 80 | 26 | 85 | 82 | 32 | 77 | 86 | 100 | 68 | 74 | 77 | 85 | 76 | 77 | 69 | 80 |
| | 92 | 39 | 75 | 44 | 73 | 79 | 79 | 74 | | | | | | | | | | | |
| 258 | 95 | 80 | 93 | 31 | 16 | 91 | 79 | 28 | 95 | 74 | 77 | 77 | 65 | 80 | 81 | 55 | 84 | 58 | 70 |
| | 75 | 23 | 56 | 28 | 56 | 51 | 74 | 58 | | | | | | | | | | | |
| 259 | 74 | 78 | 92 | 95 | 18 | 91 | 83 | -4 | 70 | 76 | 95 | 62 | 54 | 90 | 83 | 65 | 73 | 63 | 84 |
| | 76 | 27 | 56 | 40 | 60 | 54 | 72 | 56 | | | | | | | | | | | |
| 260 | 100 | 92 | 94 | 75 | 26 | 87 | 80 | 35 | 102 | 76 | 70 | 75 | 63 | 86 | 59 | 64 | 99 | 75 | 77 |
| | 90 | 39 | 64 | 50 | 71 | 80 | 77 | 73 | | | | | | | | | | | |
| 261 | 83 | 95 | 76 | 76 | 19 | 86 | 103 | 31 | 103 | 94 | 95 | 67 | 66 | 99 | 91 | 77 | 81 | 66 | 94 |
| | 99 | 47 | 70 | 46 | 74 | 69 | 81 | 67 | | | | | | | | | | | |
| 262 | 80 | 86 | 98 | 101 | 37 | 84 | 92 | 23 | 90 | 82 | 86 | 75 | 57 | 97 | 94 | 77 | 102 | 85 | 102 |
| | 98 | 47 | 69 | 59 | 75 | 81 | 93 | 79 | | | | | | | | | | | |
| 263 | 77 | 71 | 77 | 77 | 32 | 87 | 64 | 28 | 76 | 93 | 82 | 59 | 61 | 86 | 81 | 66 | 69 | 73 | 79 |
| | 84 | 48 | 58 | 42 | 68 | 67 | 72 | 57 | | | | | | | | | | | |
| 264 | 72 | 62 | 72 | 88 | 12 | 81 | 73 | 9 | 77 | 73 | 85 | 48 | 54 | 66 | 68 | 66 | 66 | 61 | 80 |
| | 58 | 24 | 48 | 25 | 61 | 65 | 57 | 62 | | | | | | | | | | | |
| 265 | 71 | 88 | 82 | 77 | -2 | 94 | 82 | 15 | 76 | 94 | 81 | 64 | 78 | 75 | 99 | 54 | 56 | 67 | 81 |
| | 92 | 27 | 58 | 38 | 70 | 63 | 72 | 52 | | | | | | | | | | | |
| 266 | 80 | 80 | 77 | 104 | 17 | 81 | 79 | 21 | 86 | 70 | 78 | 61 | 64 | 87 | 74 | 50 | 85 | 65 | 80 |
| | 78 | 35 | 72 | 44 | 77 | 66 | 84 | 62 | | | | | | | | | | | |
| 267 | 79 | 91 | 94 | 88 | | | | | | | | | | | | | | | |

DATA SIMULATIE IV
EEN ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
(ZIE SIMULATIE IV POPULATIE FAKTORSTRUKTUUR)
VOOR 27 OPERATOREN BIJ 600 ELEMENTEN

VF - VOL 6

[illegible]

DATASIMULATIE IV
EEN ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
(ZIE SIMULATIE IV POPULATIE FAKTORSTRUKTUUR)
VOOR 27 OPERATOREN BIJ 600 ELEMENTEN

VFT-VULG

[illegible]

DATASIMULATIE IV
 IEN ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
 (ZIE SIMULATIE IV POPULATIE FAKTORSTRUCTUUR)
 VOOR 27 OPERATOREN EN 500 ELEMENTEN

VERVOLG

| | | | | | | | | | | | | | | | |
|-----|----------|----------|----------|-------|----------|-------|-------------|-------------|-------|-------|-------|----|-------|-------|----|
| 316 | 91 | 80 | 30101 | 60 | 85108 | 45 | 98 | 98100 | 81 | 75 | 84 | 93 | 89105 | 72 | 95 |
| | 84 | 62 | 95 | 49 | 87 | 90109 | 77 | | | | | | | | |
| 317 | 90111 | 95 | 76 | 13 | 54 | 75 | 15 | 71 | 85 | 87 | 63 | 71 | 85 | 71 | 50 |
| | 91 | 27 | 71 | 40 | 45 | 69 | 75 | 74 | | | | | | | |
| 318 | 83 | 83114 | 88 | 15 | 83 | 93 | 38 | 81 | 56 | 92 | 68 | 67 | 50 | 51 | 70 |
| | 93 | 55 | 67 | 51 | 73 | 75 | 86 | 66 | | | | | | | |
| 319 | 89 | 79 | 69 | 86 | 22 | 73 | 81 | 41 | 98 | 87 | 72 | 79 | 77 | 80 | 78 |
| | 96 | 52 | 74 | 55 | 65 | 85 | 89 | 83 | | | | | | | |
| 320 | 93 | 52 | 60 | 83 | 28 | 57 | 85 | 27109104104 | 76 | 88 | 56 | 92 | 85101 | 25 | 65 |
| | 95 | 51 | 45 | 60 | 92 | 85103 | 87 | | | | | | | | |
| 321 | 98 | 76 | 73 | 75 | 26 | 76 | 84 | 32 | 93 | 94 | 76 | 74 | 75 | 79 | 87 |
| | 67 | 54 | 76 | 53 | 62 | 66 | 87 | 74 | | | | | | | |
| 322 | 76 | 34 | 30 | 69 | 28 | 88 | 80 | 26 | 77 | 87 | 81 | 65 | 53 | 67 | 76 |
| | 75 | 29 | 59 | 40 | 59 | 70 | 71 | 53 | | | | | | | |
| 323 | 81 | 72 | 64 | 89 | 37 | 57 | 82 | 28 | 56 | 77 | 85 | 87 | 63 | 99 | 78 |
| | 82 | 46 | 68 | 47 | 66 | 81 | 74 | 59 | | | | | | | |
| 324 | 91 | 63 | 60 | 75 | 25 | 93 | 87 | 32 | 55 | 96 | 95 | 69 | 60 | 70 | 88 |
| | 79 | 23 | 65 | 29 | 67 | 70 | 81 | 63 | | | | | | | |
| 325 | 92 | 65 | 66 | 79 | 40 | 72 | 90 | 25 | 87 | 87 | 75 | 66 | 45 | 84 | 82 |
| | 80 | 43 | 69 | 38 | 73 | 68 | 88 | 75 | | | | | | | |
| 326 | 69 | 77 | 74 | 95 | 9 | 79 | 90 | 20 | 77 | 84 | 66 | 50 | 67 | 77 | 51 |
| | 92 | 24 | 65 | 21 | 73 | 67 | 69 | 55 | | | | | | | |
| 327 | 85 | 88103 | 89 | 16 | 89 | 93 | 23 | 76 | 86 | 93 | 79 | 59 | 86 | 85 | 82 |
| | 86 | 38 | 82 | 42 | 70 | 70 | 87 | 76 | | | | | | | |
| 328 | 89108 | 91 | 86 | 11 | 89 | 92 | 32 | 96 | 88 | 76 | 68 | 69 | 69 | 62 | 64 |
| | 92 | 41 | 75 | 50 | 56 | 76 | 81 | 61 | | | | | | | |
| 329 | 78 | 79 | 83103 | 26102 | 93 | 4 | 94 | 76 | 93 | 74 | 73108 | 79 | 72 | 81 | 67 |
| | 82 | 47 | 76 | 40 | 59 | 85 | 89 | 65 | | | | | | | |
| 330 | 101 | 91109107 | 37 | 34104 | 30115114 | 98 | 85 | 83108105 | 77106 | 97103 | | | | | |
| | 92 | 47 | 90 | 61 | 88 | 81104 | 80 | | | | | | | | |
| 341 | 71 | 69 | 82 | 59 | 44 | 76 | 79 | 37 | 78 | 89 | 75 | 71 | 66 | 73 | 77 |
| | 65 | 40 | 74 | 46 | 66 | 75 | 70 | 75 | | | | | | | |
| 342 | 106 | 93 | 92 | 87 | 20 | 98 | 95 | 38 | 95 | 92 | 87 | 75 | 75 | 80 | 98 |
| | 97 | 61 | 89 | 49 | 69 | 80 | 98 | 74 | | | | | | | |
| 343 | 76 | 86 | 91 | 74 | 19 | 82 | 65 | 14 | 73 | 58 | 77 | 55 | 58 | 68 | 72 |
| | 67 | 25 | 62 | 41 | 53 | 59 | 69 | 59 | | | | | | | |
| 344 | 67 | 73 | 94 | 77 | 28 | 78 | 76 | 4 | 69 | 84 | 94 | 60 | 57 | 68 | 81 |
| | 87 | 21 | 56 | 31 | 66 | 58 | 76 | 48 | | | | | | | |
| 345 | 84 | 86 | 83 | 92 | 18 | 67 | 96 | 9 | 84 | 76 | 79 | 56 | 56 | 84 | 86 |
| | 82 | 38 | 65 | 32 | 55 | 61 | 78 | 64 | | | | | | | |
| 346 | 72 | 79 | 77 | 71 | 14 | 75 | 91 | 32 | 93 | 85 | 84 | 62 | 70 | 71 | 88 |
| | 71 | 28 | 58 | 29 | 71 | 54 | 76 | 62 | | | | | | | |
| 347 | 85 | 90 | 81 | 84 | 25 | 77 | 95 | 23 | 77102 | 86 | 72 | 70 | 82 | 86 | 72 |
| | 85 | 30 | 62 | 39 | 80 | 70 | 86 | 64 | | | | | | | |
| 348 | 10910810 | 99 | 34102110 | 21110 | 99105 | 86 | 97109109 | 93118 | 87106 | | | | | | |
| | 107 | 58 | 90 | 72 | 98106114 | 76 | | | | | | | | | |
| 349 | 82 | 34 | 64 | 82 | 31 | 62 | 97 | 28 | 82 | 82 | 81 | 61 | 55 | 78 | 83 |
| | 76 | 42 | 68 | 45 | 67 | 69 | 88 | 68 | | | | | | | |
| 350 | 80 | 98104 | 78 | 22 | 92 | 92 | 35102103110 | 89 | 77 | 93 | 84 | 70 | 85 | 87105 | |
| | 99 | 56 | 78 | 64 | 85 | 76 | 88 | 91 | | | | | | | |

DATASIMULATIE IV
EEN ALGEMEEN FAKTOR EN VIJF SUBFAKTOREN
(ZIE SIMULATIE IV POPULATIE FAKTORSTRUKTUUR)
VOOR 27 OPERATOREN BIJ 600 ELEMENTEN

VERVOLG

| | | | | | | | | | | | | | | | | | | | |
|-----|----|-----|----|----|----|-----|-----|----|-----|-----|-----|----|----|----|----|----|-----|----|----|
| 351 | 88 | 92 | 80 | 79 | 12 | 33 | 85 | 32 | 96 | 76 | 87 | 62 | 56 | 88 | 76 | 73 | 77 | 61 | 86 |
| | 89 | 31 | 65 | 46 | 75 | 76 | 71 | 72 | | | | | | | | | | | |
| 352 | 93 | 90 | 87 | 96 | 37 | 96 | 78 | 24 | 81 | 96 | 99 | 73 | 78 | 75 | 94 | 67 | 86 | 74 | 96 |
| | 80 | 47 | 72 | 53 | 63 | 87 | 93 | 68 | | | | | | | | | | | |
| 353 | 92 | 75 | 84 | 99 | 43 | 96 | 104 | 10 | 94 | 96 | 82 | 62 | 87 | 99 | 95 | 68 | 88 | 86 | 95 |
| | 86 | 59 | 73 | 47 | 87 | 73 | 97 | 80 | | | | | | | | | | | |
| 354 | 99 | 89 | 77 | 61 | 19 | 80 | 85 | 47 | 103 | 71 | 81 | 68 | 64 | 80 | 84 | 69 | 90 | 65 | 77 |
| | 91 | 48 | 70 | 52 | 73 | 63 | 90 | 69 | | | | | | | | | | | |
| 355 | 83 | 85 | 91 | 69 | 23 | 99 | 83 | 40 | 89 | 84 | 92 | 79 | 75 | 84 | 86 | 73 | 84 | 85 | 71 |
| | 84 | 27 | 77 | 48 | 77 | 66 | 79 | 61 | | | | | | | | | | | |
| 356 | 94 | 99 | 95 | 97 | 15 | 86 | 85 | 14 | 97 | 85 | 79 | 56 | 76 | 81 | 95 | 68 | 88 | 67 | 74 |
| | 85 | 51 | 67 | 45 | 72 | 63 | 83 | 66 | | | | | | | | | | | |
| 357 | 95 | 63 | 87 | 86 | 23 | 83 | 97 | 28 | 97 | 87 | 90 | 78 | 64 | 84 | 81 | 66 | 90 | 65 | 85 |
| | 89 | 33 | 67 | 34 | 71 | 73 | 81 | 78 | | | | | | | | | | | |
| 358 | 93 | 77 | 79 | 76 | 23 | 102 | 93 | 39 | 82 | 89 | 105 | 60 | 64 | 79 | 94 | 67 | 86 | 78 | 76 |
| | 78 | 37 | 74 | 34 | 72 | 69 | 85 | 85 | | | | | | | | | | | |
| 359 | 88 | 101 | 71 | 97 | 13 | 81 | 102 | 18 | 100 | 109 | 78 | 92 | 74 | 98 | 98 | 77 | 98 | 79 | 96 |
| | 92 | 64 | 86 | 55 | 78 | 90 | 100 | 88 | | | | | | | | | | | |
| 360 | 88 | 93 | 82 | 86 | 30 | 100 | 80 | 29 | 83 | 98 | 81 | 68 | 76 | 94 | 93 | 78 | 94 | 73 | 85 |
| | 79 | 47 | 73 | 44 | 84 | 80 | 86 | 61 | | | | | | | | | | | |
| 361 | 96 | 78 | 65 | 83 | 26 | 58 | 95 | 21 | 81 | 78 | 77 | 56 | 56 | 82 | 81 | 76 | 84 | 67 | 86 |
| | 82 | 49 | 65 | 41 | 58 | 71 | 89 | 73 | | | | | | | | | | | |
| 362 | 86 | 82 | 84 | 87 | 32 | 68 | 71 | 22 | 67 | 80 | 83 | 75 | 63 | 74 | 81 | 59 | 74 | 76 | 86 |
| | 83 | 41 | 66 | 59 | 72 | 70 | 85 | 73 | | | | | | | | | | | |
| 363 | 67 | 83 | 85 | 90 | 34 | 96 | 85 | 24 | 88 | 94 | 88 | 58 | 75 | 93 | 92 | 72 | 85 | 80 | 88 |
| | 93 | 44 | 60 | 44 | 76 | 67 | 78 | 63 | | | | | | | | | | | |
| 364 | 80 | 95 | 81 | 74 | 16 | 85 | 74 | 23 | 100 | 65 | 93 | 63 | 75 | 75 | 80 | 78 | 87 | 75 | 85 |
| | 82 | 44 | 72 | 41 | 75 | 70 | 81 | 76 | | | | | | | | | | | |
| 365 | 87 | 79 | 95 | 84 | 38 | 51 | 93 | 43 | 90 | 94 | 86 | 79 | 78 | 88 | 91 | 85 | 90 | 70 | 84 |
| | 73 | 42 | 79 | 43 | 80 | 62 | 84 | 72 | | | | | | | | | | | |
| 366 | 87 | 83 | 81 | 84 | 24 | 78 | 91 | 28 | 86 | 72 | 77 | 69 | 53 | 64 | 91 | 73 | 104 | 68 | 86 |
| | 89 | 44 | 74 | 37 | 81 | 56 | 91 | 67 | | | | | | | | | | | |
| 367 | 90 | 83 | 95 | 8 | | | | | | | | | | | | | | | |

DATASIMULATIE IV
EEN ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
(ZIE SIMULATIE IV POPULATIE FAKTORSTRUKTUUR)
VUUR 27 OPERATOREN BIJ 600 ELEMENTEN

VERVOLG

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|----|-----|------|----|-----|-----|----|----|----|-----|----|----|----|-----|----|----|----|-----|
| 401 | 64 | 64 | 89 | 79 | 16 | 80 | 79 | 26 | 73 | 92 | 86 | 48 | 59 | 88 | 78 | 69 | 79 | 61 | 76 |
| | 76 | 33 | 50 | 39 | 57 | 67 | 69 | 77 | | | | | | | | | | | |
| 402 | 87 | 89 | 85 | 73 | 31 | 01 | 71 | 43 | 77 | 63 | 86 | 71 | 65 | 77 | 76 | 76 | 92 | 74 | 77 |
| | 100 | 35 | 74 | 56 | 82 | 71 | 81 | 71 | | | | | | | | | | | |
| 403 | 92 | 95 | 85 | 76 | 22 | 84 | 98 | 29 | 86 | 80 | 88 | 76 | 69 | 80 | 85 | 64 | 78 | 71 | 79 |
| | 87 | 41 | 68 | 44 | 70 | 71 | 85 | 65 | | | | | | | | | | | |
| 404 | 84 | 83 | 90 | 97 | 32 | 89 | 93 | 19 | 95 | 97 | 93 | 60 | 81 | 83 | 99 | 62 | 91 | 74 | 85 |
| | 93 | 41 | 76 | 41 | 79 | 72 | 89 | 71 | | | | | | | | | | | |
| 405 | 97 | 89 | 101 | 99 | 33 | 108 | 102 | 30 | 96 | 89 | 99 | 74 | 83 | 95 | 103 | 77 | 85 | 71 | 90 |
| | 91 | 46 | 89 | 48 | 73 | 76 | 87 | 69 | | | | | | | | | | | |
| 406 | 89 | 94 | 77 | 72 | 28 | 75 | 90 | 38 | 91 | 89 | 83 | 64 | 76 | 85 | 85 | 79 | 87 | 67 | 80 |
| | 89 | 39 | 84 | 45 | 70 | 72 | 89 | 71 | | | | | | | | | | | |
| 407 | 89 | 76 | 101 | 92 | 41 | 88 | 98 | 27 | 85 | 93 | 96 | 58 | 71 | 90 | 78 | 76 | 94 | 76 | 104 |
| | 86 | 41 | 81 | 50 | 66 | 70 | 86 | 85 | | | | | | | | | | | |
| 408 | 87 | 67 | 88 | 92 | 30 | 84 | 83 | 30 | 89 | 74 | 69 | 68 | 62 | 76 | 73 | 65 | 91 | 62 | 79 |
| | 68 | 36 | 64 | 40 | 58 | 64 | 74 | 57 | | | | | | | | | | | |
| 409 | 76 | 84 | 90 | 80 | 28 | 87 | 93 | 34 | 96 | 88 | 98 | 85 | 55 | 82 | 85 | 79 | 83 | 58 | 84 |
| | 79 | 30 | 77 | 44 | 85 | 81 | 85 | 68 | | | | | | | | | | | |
| 410 | 84 | 92 | 75 | 67 | 13 | 78 | 67 | 34 | 86 | 75 | 85 | 57 | 82 | 70 | 62 | 61 | 81 | 43 | 65 |
| | 77 | 32 | 46 | 43 | 63 | 67 | 80 | 60 | | | | | | | | | | | |
| 411 | 89 | 84 | 97 | 73 | 15 | 80 | 81 | 25 | 91 | 85 | 92 | 80 | 64 | 75 | 86 | 73 | 93 | 70 | 81 |
| | 87 | 41 | 74 | 42 | 64 | 59 | 83 | 68 | | | | | | | | | | | |
| 412 | 86 | 89 | 75 | 95 | 30 | 81 | 81 | 35 | 82 | 84 | 84 | 66 | 80 | 80 | 91 | 61 | 89 | 68 | 96 |
| | 96 | 53 | 78 | 65 | 79 | 73 | 86 | 73 | | | | | | | | | | | |
| 413 | 83 | 80 | 94 | 74 | 28 | 87 | 91 | 28 | 88 | 94 | 101 | 79 | 63 | 79 | 97 | 67 | 89 | 82 | 90 |
| | 86 | 47 | 63 | 47 | 77 | 70 | 85 | 60 | | | | | | | | | | | |
| 414 | 75 | 91 | 88 | 70 | 23 | 84 | 79 | 35 | 71 | 95 | 89 | 65 | 68 | 83 | 77 | 73 | 82 | 75 | 81 |
| | 97 | 29 | 74 | 42 | 78 | 72 | 77 | 66 | | | | | | | | | | | |
| 415 | 85 | 82 | 89 | 80 | 17 | 85 | 92 | 12 | 96 | 81 | 80 | 67 | 67 | 87 | 82 | 70 | 92 | 72 | 100 |
| | 93 | 40 | 74 | 31 | 78 | 69 | 86 | 81 | | | | | | | | | | | |
| 416 | 67 | 78 | 77 | 94 | 14 | 82 | 87 | 31 | 93 | 94 | 74 | 59 | 75 | 80 | 100 | 74 | 83 | 79 | 88 |
| | 92 | 47 | 74 | 33 | 74 | 73 | 78 | 84 | | | | | | | | | | | |
| 417 | 74 | 86 | 95 | 67</ | | | | | | | | | | | | | | | |

**DATASIMULATIE IV
EEN ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
(ZIE SIMULATIE IV POPULATIE FAKTORSTRUKTUUR)
VOOR 27 OPERATOREN BIJ 600 ELEMENTEN**

VERVOLG

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|----|----|----|-----|----|-----|----|-----|----|----|----|-----|----|-----|----|----|
| 426 | 96 | 79 | 94 | 90 | 33 | 86 | 96 | 20 | 85 | 85 | 100 | 77 | 81 | 94 | 76 | 78 | 100 | 85 | 91 |
| | 97 | 53 | 84 | 61 | 70 | 78 | 96 | 83 | | | | | | | | | | | |
| 427 | 91 | 92 | 86 | 75 | 10 | 87 | 91 | 33 | 88 | 86 | 86 | 55 | 68 | 89 | 84 | 86 | 85 | 70 | 86 |
| | 99 | 45 | 74 | 40 | 82 | 78 | 79 | 72 | | | | | | | | | | | |
| 428 | 88 | 96 | 82 | 93 | 15 | 91 | 91 | 33 | 89 | 89 | 93 | 68 | 70 | 82 | 94 | 73 | 95 | 78 | 93 |
| | 103 | 43 | 76 | 46 | 69 | 79 | 100 | 81 | | | | | | | | | | | |
| 429 | 72 | 81 | 72 | 76 | 24 | 76 | 95 | 30 | 81 | 80 | 72 | 61 | 79 | 85 | 84 | 49 | 76 | 67 | 69 |
| | 75 | 33 | 57 | 31 | 66 | 66 | 76 | 57 | | | | | | | | | | | |
| 430 | 77 | 80 | 109 | 74 | 32 | 89 | 84 | 23 | 88 | 87 | 91 | 79 | 69 | 82 | 84 | 71 | 85 | 70 | 98 |
| | 79 | 32 | 78 | 45 | 72 | 84 | 91 | 71 | | | | | | | | | | | |
| 431 | 89 | 80 | 65 | 93 | 16 | 84 | 89 | 28 | 96 | 82 | 80 | 52 | 55 | 82 | 81 | 77 | 94 | 73 | 88 |
| | 82 | 38 | 64 | 45 | 77 | 49 | 72 | 50 | | | | | | | | | | | |
| 432 | 94 | 88 | 85 | 86 | 30 | 91 | 99 | 26 | 79 | 94 | 108 | 75 | 65 | 90 | 103 | 72 | 83 | 74 | 97 |
| | 85 | 47 | 71 | 51 | 81 | 75 | 87 | 78 | | | | | | | | | | | |
| 433 | 95 | 85 | 90 | 78 | 19 | 81 | 76 | 26 | 99 | 86 | 95 | 62 | 81 | 84 | 85 | 87 | 85 | 77 | 89 |
| | 89 | 41 | 79 | 47 | 75 | 71 | 92 | 75 | | | | | | | | | | | |
| 434 | 86 | 64 | 88 | 66 | 14 | 74 | 61 | 36 | 88 | 92 | 76 | 87 | 70 | 71 | 90 | 51 | 81 | 64 | 72 |
| | 89 | 38 | 64 | 39 | 65 | 65 | 78 | 69 | | | | | | | | | | | |
| 435 | 83 | 105 | 91 | 98 | 45 | 78 | 114 | 38 | 94 | 97 | 94 | 75 | 94 | 86 | 109 | 81 | 97 | 74 | 96 |
| | 91 | 61 | 77 | 49 | 91 | 78 | 102 | 76 | | | | | | | | | | | |
| 436 | 92 | 75 | 80 | 93 | 39 | 88 | 91 | 31 | 106 | 84 | 106 | 65 | 83 | 91 | 89 | 81 | 91 | 85 | 95 |
| | 36 | 48 | 94 | 59 | 83 | 84 | 94 | 80 | | | | | | | | | | | |
| 437 | 82 | 84 | 87 | 76 | 12 | 77 | 92 | 27 | 68 | 85 | 68 | 52 | 58 | 81 | 89 | 61 | 68 | 71 | 72 |
| | 73 | 21 | 43 | 36 | 65 | 64 | 73 | 44 | | | | | | | | | | | |
| 438 | 96 | 89 | 80 | 89 | 24 | 81 | 82 | 19 | 82 | 86 | 95 | 66 | 70 | 84 | 90 | 78 | 82 | 55 | 89 |
| | 85 | 45 | 80 | 40 | 62 | 65 | 71 | 68 | | | | | | | | | | | |
| 439 | 80 | 85 | 86 | 72 | 16 | 77 | 72 | 29 | 63 | 76 | 81 | 62 | 68 | 86 | 72 | 77 | 74 | 52 | 77 |
| | 92 | 42 | 62 | 47 | 50 | 64 | 71 | 66 | | | | | | | | | | | |
| 440 | 90 | 77 | 65 | 71 | 25 | 83 | 79 | 38 | 88 | 81 | 73 | 66 | 75 | 86 | 93 | 52 | 80 | 46 | 76 |
| | 75 | 42 | 71 | 30 | 69 | 65 | 75 | 68 | | | | | | | | | | | |
| 441 | 84 | 80 | 97 | 86 | 33 | 96 | 86 | 15 | 99 | 80 | 104 | 66 | 91 | 86 | 80 | 78 | 95 | 63 | 99 |
| | 79 | 45 | 76 | 52 | 81 | 77 | 86 | 83 | | | | | | | | | | | |
| 442 | 82 | 84 | 86 | | | | | | | | | | | | | | | | |

DATASIMULATIE IV
EEN ALGEMENE FACTOR EN VIJF SUBFACTOREN
(ZIE SIMULATIE IV POPULATIE FACTORSTRUCTUUR)
VOOR 27 OPERATOREN BIJ 600 ELEMENTEN

VERVOLG

[illegible]

DATASIMULATIE IV
EEN ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
(ZIE SIMULATIE IV POPULATIE FAKTORSTRUKTUUR)
VOOR 27 OPERATOREN BIJ 600 ELEMENTEN

VERVOLG

[illegible]

DATASIMULATIE IV
EEN ALGEMEEN FAKTOR EN VIJF SUBFAKTOREN
(ZIE SIMULATIE IV POPULATIE FAKTORSTRUKTUUR)
VOOR 27 OPERATOREN BIJ 600 ELEMENTEN

VERVOLG

[illegible]

DATASIMULATIE IV
 EEN ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
 (ZIE SIMULATIE IV POPULATIE FAKTORSTRUKTUUR)
 VOOR 27 OPERATOREN BIJ 600 ELEMENTEN

VFR VOLG

| | | | | | | | | | | | | | | | | | | | |
|-----|----|----|-----|----|------|----|-----|----|-----|----|----|----|----|-----|-----|----|----|----|----|
| 525 | 71 | 31 | 99 | 79 | 26 | 82 | 70 | 45 | 72 | 67 | 97 | 59 | 67 | 72 | 83 | 71 | 70 | 80 | 79 |
| | 38 | 25 | 54 | 49 | 71 | 70 | 73 | 67 | | | | | | | | | | | |
| 527 | 77 | 38 | 71 | 80 | 17 | 79 | 85 | 30 | 99 | 87 | 87 | 76 | 68 | 95 | 93 | 79 | 98 | 76 | 86 |
| | 74 | 51 | 77 | 43 | 69 | 65 | 92 | 70 | | | | | | | | | | | |
| 528 | 88 | 87 | 83 | 36 | 38 | 93 | 90 | 18 | 86 | 94 | 98 | 74 | 76 | 87 | 78 | 57 | 82 | 72 | 84 |
| | 43 | 33 | 77 | 35 | 61 | 66 | 70 | 67 | | | | | | | | | | | |
| 529 | 74 | 70 | 92 | 83 | 18 | 64 | 79 | 7 | 69 | 88 | 73 | 50 | 50 | 84 | 78 | 65 | 68 | 52 | 80 |
| | 73 | 37 | 60 | 27 | 45 | 64 | 69 | 63 | | | | | | | | | | | |
| 530 | 76 | 63 | 82 | 81 | 44 | 86 | 75 | 21 | 65 | 62 | 84 | 59 | 62 | 88 | 80 | 81 | 68 | 70 | 90 |
| | 65 | 32 | 68 | 48 | 52 | 66 | 65 | 66 | | | | | | | | | | | |
| 531 | 74 | 92 | 103 | 76 | 13 | 92 | 97 | 54 | 86 | 87 | 91 | 50 | 86 | 77 | 86 | 72 | 78 | 71 | 73 |
| | 90 | 36 | 61 | 41 | 67 | 64 | 80 | 83 | | | | | | | | | | | |
| 532 | 73 | 75 | 66 | 79 | 16 | 70 | 79 | 27 | 74 | 83 | 70 | 63 | 55 | 73 | 87 | 74 | 69 | 60 | 71 |
| | 69 | 33 | 64 | 30 | 63 | 54 | 75 | 59 | | | | | | | | | | | |
| 533 | 80 | 79 | 85 | 87 | 27 | 88 | 74 | 18 | 75 | 91 | 90 | 68 | 54 | 104 | 93 | 67 | 85 | 60 | 86 |
| | 72 | 51 | 70 | 27 | 78 | 75 | 78 | 70 | | | | | | | | | | | |
| 534 | 83 | 77 | 79 | 86 | 37 | 88 | 90 | 30 | 97 | 84 | 94 | 73 | 61 | 90 | 89 | 67 | 92 | 83 | 95 |
| | 92 | 46 | 77 | 59 | 93 | 71 | 93 | 78 | | | | | | | | | | | |
| 535 | 75 | 73 | 81 | 90 | 34 | 91 | 78 | 32 | 92 | 87 | 88 | 60 | 66 | 94 | 86 | 82 | 86 | 66 | 81 |
| | 82 | 45 | 67 | 41 | 71 | 72 | 78 | 74 | | | | | | | | | | | |
| 536 | 70 | 77 | 90 | 72 | 8 | 78 | 80 | 18 | 103 | 69 | 75 | 57 | 55 | 77 | 56 | 49 | 87 | 53 | 66 |
| | 79 | 32 | 64 | 59 | 51 | 57 | 78 | 56 | | | | | | | | | | | |
| 537 | 84 | 38 | 87 | 79 | 16 | 85 | 102 | 18 | 94 | 84 | 97 | 68 | 64 | 99 | 105 | 72 | 73 | 84 | 95 |
| | 95 | 52 | 69 | 34 | 75 | 77 | 81 | 80 | | | | | | | | | | | |
| 538 | 80 | 74 | 77 | 85 | 42 | 73 | 93 | 36 | 82 | 89 | 62 | 77 | 63 | 86 | 91 | 58 | 82 | 73 | 85 |
| | 68 | 51 | 78 | 52 | 73 | 77 | 87 | 64 | | | | | | | | | | | |
| 539 | 90 | 78 | 76 | 77 | 23 | 75 | 89 | 14 | 84 | 71 | 75 | 74 | 60 | 75 | 76 | 43 | 72 | 62 | 67 |
| | 72 | 29 | 51 | 34 | 70 | 54 | 74 | 66 | | | | | | | | | | | |
| 540 | 74 | 79 | 74 | 78 | 40 | 62 | 73 | 8 | 69 | 72 | 73 | 72 | 59 | 80 | 72 | 74 | 63 | 73 | 85 |
| | 76 | 40 | 62 | 42 | 61 | 61 | 57 | 66 | | | | | | | | | | | |
| 541 | 95 | 81 | 77 | 94 | 31 | 72 | 88 | 32 | 84 | 82 | 85 | 98 | 89 | 109 | 88 | 76 | 92 | 67 | 92 |
| | 97 | 73 | 85 | 59 | 90 | 78 | 93 | 90 | | | | | | | | | | | |
| 542 | 78 | 38 | 88 | 71 | 30</ | | | | | | | | | | | | | | |

DATASIMULATIE IV
EEN ALGEMENE FAKTOR EN VIJF SUBFAKTOREN
(ZIE SIMULATIE IV POPULATIE FAKTORSTRUKTUUR)
VOOR 27 OPERATOREN BIJ 600 ELEMENTEN

VFR VOLG

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|----|-----|----|----|-----|-----|----|----|-----|----|----|----|-----|----|----|-----|----|----|
| 575 | 35 | 67 | 99 | 93 | 34 | 70 | 87 | 31 | 92 | 79 | 84 | 53 | 68 | 92 | 83 | 65 | 88 | 72 | 86 |
| | 76 | 30 | 76 | 41 | 70 | 69 | 88 | 67 | | | | | | | | | | | |
| 577 | 85 | 75 | 81 | 60 | 25 | 68 | 93 | 16 | 77 | 82 | 79 | 65 | 86 | 78 | 82 | 77 | 84 | 66 | 91 |
| | 64 | 42 | 67 | 40 | 60 | 69 | 80 | 66 | | | | | | | | | | | |
| 578 | 105 | 81 | 89 | 74 | 27 | 75 | 99 | 18 | 97 | 89 | 67 | 86 | 68 | 81 | 93 | 76 | 101 | 66 | 81 |
| | 76 | 50 | 78 | 37 | 64 | 65 | 83 | 67 | | | | | | | | | | | |
| 579 | 76 | 79 | 106 | 86 | 33 | 89 | 80 | 27 | 74 | 70 | 98 | 70 | 65 | 77 | 85 | 58 | 72 | 67 | 83 |
| | 60 | 34 | 66 | 32 | 69 | 57 | 75 | 58 | | | | | | | | | | | |
| 580 | 77 | 76 | 73 | 87 | 12 | 78 | 83 | 10 | 79 | 83 | 90 | 70 | 61 | 93 | 77 | 69 | 85 | 69 | 70 |
| | 81 | 27 | 64 | 33 | 63 | 63 | 82 | 67 | | | | | | | | | | | |
| 581 | 91 | 73 | 84 | 86 | 16 | 70 | 69 | 21 | 73 | 99 | 83 | 69 | 57 | 91 | 66 | 66 | 86 | 71 | 82 |
| | 85 | 44 | 78 | 42 | 69 | 65 | 78 | 63 | | | | | | | | | | | |
| 582 | 87 | 83 | 91 | 92 | 29 | 108 | 99 | 25 | 87 | 94 | 87 | 77 | 97 | 85 | 85 | 75 | 105 | 79 | 96 |
| | 88 | 42 | 80 | 46 | 81 | 79 | 95 | 78 | | | | | | | | | | | |
| 583 | 95 | 95 | 74 | 75 | 31 | 74 | 86 | 29 | 88 | 94 | 80 | 70 | 61 | 91 | 78 | 80 | 93 | 74 | 76 |
| | 77 | 41 | 73 | 52 | 70 | 80 | 87 | 62 | | | | | | | | | | | |
| 584 | 75 | 35 | 80 | 35 | 2 | 69 | 72 | 23 | 73 | 82 | 74 | 51 | 62 | 69 | 80 | 47 | 83 | 50 | 69 |
| | 78 | 33 | 54 | 42 | 57 | 50 | 67 | 55 | | | | | | | | | | | |
| 585 | 77 | 81 | 83 | 95 | 32 | 86 | 106 | 34 | 91 | 100 | 70 | 68 | 75 | 90 | 93 | 65 | 80 | 71 | 89 |
| | 81 | 55 | 87 | 55 | 82 | 55 | 78 | 80 | | | | | | | | | | | |
| 586 | 72 | 72 | 84 | 80 | 39 | 90 | 79 | 8 | 71 | 91 | 78 | 70 | 70 | 90 | 83 | 65 | 70 | 66 | 85 |
| | 67 | 20 | 75 | 27 | 75 | 57 | 71 | 64 | | | | | | | | | | | |
| 587 | 89 | 89 | 83 | 94 | 34 | 73 | 79 | 31 | 98 | 89 | 86 | 82 | 67 | 86 | 87 | 70 | 90 | 78 | 96 |
| | 88 | 50 | 74 | 58 | 63 | 72 | 85 | 75 | | | | | | | | | | | |
| 588 | 88 | 91 | 94 | 82 | 47 | 95 | 82 | 37 | 89 | 75 | 91 | 75 | 67 | 102 | 91 | 66 | 93 | 75 | 89 |
| | 97 | 43 | 62 | 57 | 80 | 91 | 96 | 87 | | | | | | | | | | | |
| 589 | 80 | 87 | 94 | 79 | 29 | 85 | 82 | 24 | 89 | 69 | 77 | 76 | 77 | 88 | 80 | 67 | 76 | 60 | 71 |
| | 72 | 46 | 66 | 53 | 68 | 62 | 76 | 70 | | | | | | | | | | | |
| 590 | 77 | 73 | 94 | 82 | 23 | 85 | 82 | 23 | 79 | 81 | 83 | 82 | 66 | 80 | 90 | 73 | 81 | 71 | 84 |
| | 78 | 34 | 81 | 36 | 52 | 62 | 80 | 69 | | | | | | | | | | | |
| 591 | 61 | 94 | 83 | 95 | 19 | 78 | 100 | 7 | 80 | 101 | 86 | 63 | 59 | 96 | 79 | 72 | 84 | 60 | 90 |
| | 86 | 39 | 67 | 35 | 65 | 62 | 68 | 68 | | | | | | | | | | | |
| 592 | 75 | 92 | 83 | 74 | | | | | | | | | | | | | | | |

DATASIMULATIE V
TWEDE GELIJKE GROEPSFAKTOREN (ZIE TABEL 6.2.2.)
VOOR 12 OPERATOREN BIJ 100 ELEMENTEN

| | | | | | | | | | | | | |
|----|-----|----|----|----|----|----|----|----|-----|----|----|----|
| 1 | 80 | 69 | 80 | 77 | 89 | 80 | 79 | 61 | 77 | 65 | 61 | 67 |
| 2 | 85 | 63 | 85 | 64 | 71 | 75 | 74 | 77 | 65 | 83 | 63 | 60 |
| 3 | 97 | 69 | 92 | 88 | 75 | 87 | 80 | 88 | 86 | 70 | 81 | 68 |
| 4 | 75 | 59 | 77 | 71 | 76 | 78 | 76 | 67 | 82 | 74 | 85 | 74 |
| 5 | 64 | 89 | 87 | 77 | 75 | 71 | 71 | 67 | 75 | 74 | 70 | 71 |
| 6 | 68 | 79 | 63 | 72 | 67 | 65 | 69 | 66 | 56 | 64 | 82 | 62 |
| 7 | 72 | 90 | 93 | 74 | 87 | 76 | 58 | 89 | 66 | 71 | 64 | 76 |
| 8 | 67 | 70 | 86 | 74 | 90 | 76 | 79 | 81 | 79 | 66 | 80 | 73 |
| 9 | 59 | 55 | 74 | 50 | 72 | 48 | 80 | 73 | 75 | 64 | 63 | 57 |
| 10 | 78 | 84 | 92 | 66 | 83 | 80 | 71 | 86 | 88 | 61 | 75 | 74 |
| 11 | 75 | 73 | 81 | 75 | 82 | 76 | 73 | 72 | 71 | 70 | 89 | 79 |
| 12 | 96 | 82 | 79 | 76 | 78 | 85 | 70 | 56 | 75 | 73 | 66 | 67 |
| 13 | 69 | 71 | 75 | 64 | 69 | 65 | 60 | 77 | 74 | 68 | 63 | 58 |
| 14 | 78 | 63 | 71 | 47 | 74 | 61 | 79 | 63 | 78 | 69 | 69 | 84 |
| 15 | 65 | 60 | 63 | 65 | 79 | 74 | 82 | 74 | 74 | 69 | 80 | 81 |
| 16 | 109 | 72 | 84 | 77 | 74 | 75 | 69 | 81 | 84 | 82 | 82 | 74 |
| 17 | 80 | 72 | 85 | 64 | 85 | 88 | 94 | 90 | 87 | 79 | 75 | 82 |
| 18 | 55 | 73 | 78 | 55 | 75 | 65 | 84 | 75 | 62 | 71 | 66 | 64 |
| 19 | 65 | 71 | 59 | 74 | 80 | 50 | 62 | 63 | 70 | 61 | 80 | 66 |
| 20 | 73 | 70 | 70 | 73 | 93 | 76 | 80 | 79 | 75 | 82 | 75 | 81 |
| 21 | 75 | 68 | 65 | 68 | 69 | 70 | 76 | 72 | 75 | 58 | 62 | 73 |
| 22 | 93 | 75 | 67 | 69 | 97 | 72 | 59 | 42 | 64 | 59 | 68 | 67 |
| 23 | 84 | 54 | 75 | 85 | 77 | 66 | 61 | 57 | 70 | 67 | 72 | 60 |
| 24 | 67 | 59 | 68 | 57 | 80 | 50 | 66 | 72 | 76 | 55 | 87 | 66 |
| 25 | 68 | 59 | 61 | 67 | 68 | 66 | 66 | 63 | 66 | 78 | 71 | 70 |
| 26 | 65 | 57 | 80 | 51 | 62 | 67 | 75 | 56 | 67 | 60 | 78 | 78 |
| 27 | 66 | 57 | 85 | 64 | 62 | 70 | 81 | 75 | 58 | 76 | 65 | 67 |
| 28 | 85 | 76 | 74 | 82 | 84 | 85 | 83 | 64 | 82 | 80 | 73 | 68 |
| 29 | 74 | 40 | 64 | 66 | 73 | 64 | 90 | 78 | 84 | 60 | 71 | 74 |
| 30 | 86 | 86 | 87 | 75 | 85 | 89 | 71 | 74 | 75 | 59 | 64 | 72 |
| 31 | 76 | 68 | 78 | 61 | 59 | 72 | 83 | 74 | 81 | 65 | 70 | 66 |
| 32 | 79 | 67 | 80 | 68 | 80 | 71 | 76 | 72 | 99 | 73 | 63 | 81 |
| 33 | 49 | 58 | 65 | 64 | 56 | 54 | 75 | 81 | 66 | 71 | 61 | 72 |
| 34 | 83 | 65 | 62 | 72 | 69 | 65 | 77 | 80 | 66 | 57 | 81 | 67 |
| 35 | 92 | 94 | 66 | 72 | 79 | 75 | 74 | 71 | 58 | 76 | 83 | 71 |
| 36 | 79 | 49 | 69 | 60 | 76 | 48 | 69 | 60 | 66 | 74 | 88 | 79 |
| 37 | 95 | 80 | 74 | 64 | 79 | 81 | 84 | 71 | 63 | 60 | 88 | 79 |
| 38 | 86 | 46 | 69 | 58 | 76 | 52 | 66 | 65 | 85 | 63 | 70 | 61 |
| 39 | 58 | 68 | 70 | 67 | 75 | 58 | 72 | 71 | 75 | 55 | 61 | 72 |
| 40 | 61 | 86 | 76 | 76 | 87 | 63 | 73 | 79 | 79 | 66 | 83 | 79 |
| 41 | 85 | 73 | 97 | 76 | 87 | 80 | 89 | 87 | 78 | 64 | 59 | 77 |
| 42 | 86 | 68 | 81 | 81 | 81 | 80 | 77 | 62 | 103 | 76 | 94 | 65 |
| 43 | 64 | 64 | 75 | 76 | 86 | 69 | 70 | 49 | 70 | 60 | 77 | 85 |
| 44 | 73 | 67 | 66 | 76 | 71 | 73 | 57 | 68 | 60 | 74 | 69 | 61 |
| 45 | 73 | 51 | 77 | 56 | 77 | 71 | 64 | 53 | 67 | 68 | 62 | 65 |
| 46 | 87 | 87 | 91 | 71 | 87 | 80 | 59 | 70 | 61 | 76 | 77 | 68 |
| 47 | 84 | 66 | 85 | 73 | 82 | 65 | 73 | 48 | 82 | 60 | 61 | 58 |
| 48 | 78 | 73 | 77 | 79 | 75 | 62 | 62 | 72 | 99 | 90 | 83 | 68 |
| 49 | 69 | 53 | 58 | 69 | 61 | 68 | 80 | 55 | 69 | 72 | 77 | 57 |
| 50 | 61 | 55 | 71 | 59 | 68 | 60 | 65 | 77 | 77 | 68 | 65 | 61 |

DATASIMULATIE V
 TWE E GELIJKE GROEPSFAKTOREN (ZIE TABEL 6.2.2.)
 VOOR 12 OPERATOREN BIJ 100 ELEMENTEN

VERVOLG

| | | | | | | | | | | | | |
|-----|-----|----|----|----|----|----|----|----|-----|----|----|----|
| 51 | 68 | 54 | 62 | 57 | 50 | 40 | 63 | 61 | 73 | 46 | 62 | 66 |
| 52 | 80 | 76 | 78 | 75 | 76 | 73 | 91 | 71 | 76 | 53 | 85 | 70 |
| 53 | 73 | 60 | 76 | 71 | 84 | 69 | 59 | 64 | 68 | 64 | 70 | 54 |
| 54 | 81 | 56 | 88 | 84 | 72 | 83 | 74 | 65 | 91 | 88 | 70 | 78 |
| 55 | 71 | 82 | 70 | 74 | 85 | 71 | 77 | 66 | 67 | 57 | 60 | 61 |
| 56 | 48 | 59 | 77 | 57 | 67 | 76 | 75 | 68 | 72 | 50 | 79 | 59 |
| 57 | 74 | 71 | 77 | 67 | 67 | 72 | 78 | 86 | 62 | 76 | 83 | 82 |
| 58 | 80 | 70 | 79 | 57 | 94 | 77 | 82 | 67 | 81 | 72 | 86 | 73 |
| 59 | 62 | 73 | 89 | 45 | 69 | 60 | 72 | 69 | 80 | 82 | 82 | 79 |
| 60 | 79 | 47 | 63 | 66 | 75 | 68 | 69 | 67 | 63 | 53 | 63 | 69 |
| 61 | 64 | 63 | 72 | 72 | 55 | 70 | 75 | 49 | 73 | 71 | 79 | 79 |
| 62 | 71 | 76 | 81 | 62 | 79 | 67 | 64 | 66 | 66 | 73 | 73 | 70 |
| 63 | 56 | 47 | 70 | 50 | 50 | 51 | 72 | 78 | 77 | 76 | 60 | 67 |
| 64 | 86 | 71 | 71 | 64 | 76 | 76 | 73 | 56 | 50 | 56 | 54 | 66 |
| 65 | 84 | 82 | 72 | 95 | 97 | 86 | 80 | 64 | 84 | 63 | 71 | 79 |
| 66 | 66 | 81 | 69 | 71 | 76 | 62 | 72 | 77 | 67 | 46 | 67 | 64 |
| 67 | 87 | 75 | 74 | 50 | 84 | 74 | 75 | 76 | 94 | 66 | 69 | 61 |
| 68 | 71 | 56 | 64 | 87 | 64 | 68 | 72 | 70 | 88 | 78 | 61 | 68 |
| 69 | 63 | 65 | 78 | 57 | 68 | 74 | 80 | 74 | 82 | 40 | 80 | 57 |
| 70 | 66 | 63 | 68 | 63 | 71 | 68 | 56 | 71 | 70 | 57 | 58 | 65 |
| 71 | 100 | 96 | 95 | 88 | 78 | 91 | 96 | 74 | 72 | 89 | 77 | 69 |
| 72 | 63 | 67 | 58 | 55 | 75 | 52 | 83 | 67 | 62 | 69 | 70 | 58 |
| 73 | 77 | 54 | 65 | 60 | 75 | 70 | 79 | 86 | 88 | 79 | 92 | 78 |
| 74 | 56 | 73 | 78 | 60 | 57 | 65 | 70 | 78 | 84 | 57 | 68 | 64 |
| 75 | 72 | 72 | 91 | 72 | 63 | 78 | 77 | 80 | 74 | 33 | 74 | 78 |
| 76 | 78 | 51 | 59 | 71 | 78 | 69 | 74 | 64 | 82 | 80 | 63 | 66 |
| 77 | 93 | 85 | 73 | 72 | 76 | 75 | 81 | 64 | 76 | 73 | 89 | 83 |
| 78 | 84 | 68 | 74 | 65 | 89 | 74 | 75 | 66 | 77 | 58 | 73 | 80 |
| 79 | 103 | 79 | 86 | 87 | 92 | 84 | 93 | 89 | 90 | 88 | 79 | 73 |
| 80 | 90 | 68 | 87 | 71 | 80 | 83 | 77 | 61 | 81 | 59 | 73 | 64 |
| 81 | 73 | 67 | 82 | 51 | 75 | 78 | 66 | 64 | 78 | 52 | 72 | 55 |
| 82 | 70 | 69 | 66 | 72 | 73 | 78 | 73 | 68 | 65 | 58 | 77 | 76 |
| 83 | 78 | 76 | 71 | 84 | 81 | 64 | 76 | 62 | 79 | 64 | 79 | 76 |
| 84 | 77 | 79 | 78 | 80 | 80 | 72 | 73 | 57 | 72 | 83 | 57 | 58 |
| 85 | 71 | 69 | 91 | 69 | 76 | 74 | 81 | 85 | 59 | 69 | 91 | 78 |
| 86 | 73 | 87 | 91 | 66 | 91 | 72 | 82 | 56 | 84 | 64 | 80 | 57 |
| 87 | 76 | 73 | 67 | 64 | 79 | 65 | 81 | 81 | 80 | 75 | 83 | 82 |
| 88 | 85 | 69 | 78 | 72 | 79 | 71 | 79 | 75 | 107 | 71 | 88 | 86 |
| 89 | 83 | 80 | 82 | 87 | 92 | 90 | 66 | 63 | 78 | 58 | 70 | 72 |
| 90 | 82 | 62 | 91 | 65 | 82 | 79 | 90 | 58 | 72 | 87 | 87 | 66 |
| 91 | 78 | 64 | 71 | 65 | 66 | 64 | 80 | 78 | 83 | 71 | 85 | 67 |
| 92 | 69 | 82 | 73 | 58 | 70 | 71 | 76 | 68 | 88 | 79 | 82 | 59 |
| 93 | 67 | 71 | 69 | 59 | 72 | 68 | 77 | 80 | 70 | 84 | 86 | 67 |
| 94 | 54 | 55 | 60 | 52 | 55 | 40 | 71 | 56 | 63 | 74 | 80 | 67 |
| 95 | 81 | 72 | 78 | 49 | 70 | 83 | 75 | 74 | 79 | 56 | 72 | 76 |
| 96 | 77 | 77 | 76 | 82 | 79 | 64 | 58 | 54 | 63 | 57 | 66 | 64 |
| 97 | 78 | 55 | 81 | 60 | 66 | 52 | 81 | 72 | 88 | 76 | 93 | 84 |
| 98 | 78 | 65 | 73 | 70 | 64 | 69 | 72 | 76 | 83 | 68 | 82 | 75 |
| 99 | 78 | 63 | 65 | 76 | 69 | 71 | 79 | 62 | 62 | 66 | 74 | 71 |
| 100 | 85 | 80 | 83 | 85 | 73 | 78 | 65 | 55 | 66 | 44 | 65 | 57 |

DATASIMULATIE VI
 DRIE GROEPSFAKTOREN EN EEN SUBFAKTOR
 (ZIE SIMULATIE VI POPULATIE FAKTORSTRUKTUUR)
 VOOR 14 OPERATOREN BIJ 100 ELEMENTEN

| | | | | | | | | | | | | | | |
|----|----|----|----|-----|-----|----|-----|----|-----|-----|----|----|----|----|
| 1 | 75 | 23 | 46 | 104 | 102 | 71 | 87 | 11 | 80 | 87 | 77 | 76 | 63 | 61 |
| 2 | 62 | 17 | 31 | 58 | 70 | 59 | 76 | 20 | 80 | 65 | 54 | 62 | 70 | 75 |
| 3 | 55 | 43 | 25 | 78 | 86 | 68 | 77 | 29 | 79 | 69 | 65 | 38 | 80 | 77 |
| 4 | 80 | 25 | 42 | 71 | 82 | 70 | 87 | 39 | 79 | 76 | 75 | 78 | 66 | 69 |
| 5 | 90 | 45 | 48 | 93 | 103 | 82 | 88 | 25 | 80 | 95 | 80 | 63 | 71 | 67 |
| 6 | 84 | 20 | 30 | 94 | 83 | 65 | 89 | 30 | 85 | 94 | 78 | 78 | 74 | 70 |
| 7 | 77 | 41 | 35 | 97 | 75 | 85 | 77 | 6 | 76 | 83 | 55 | 55 | 69 | 70 |
| 8 | 95 | 39 | 43 | 70 | 91 | 79 | 91 | 24 | 86 | 82 | 65 | 63 | 72 | 63 |
| 9 | 82 | 25 | 41 | 89 | 84 | 63 | 91 | 19 | 87 | 97 | 60 | 60 | 66 | 60 |
| 10 | 71 | 27 | 41 | 85 | 83 | 74 | 84 | 18 | 73 | 74 | 59 | 60 | 48 | 62 |
| 11 | 86 | 24 | 14 | 89 | 88 | 62 | 81 | 28 | 83 | 77 | 65 | 78 | 64 | 63 |
| 12 | 70 | 24 | 22 | 85 | 82 | 70 | 102 | 24 | 103 | 108 | 96 | 83 | 88 | 68 |
| 13 | 70 | 36 | 41 | 88 | 93 | 69 | 85 | 20 | 79 | 92 | 84 | 61 | 62 | 63 |
| 14 | 69 | 25 | 38 | 97 | 90 | 90 | 81 | 20 | 84 | 85 | 80 | 69 | 76 | 71 |
| 15 | 59 | 9 | 35 | 76 | 60 | 50 | 72 | 24 | 83 | 90 | 79 | 71 | 61 | 69 |
| 16 | 86 | 30 | 42 | 90 | 93 | 78 | 81 | 19 | 78 | 90 | 69 | 75 | 75 | 82 |
| 17 | 78 | 21 | 46 | 88 | 89 | 72 | 100 | 40 | 101 | 111 | 80 | 82 | 77 | 60 |
| 18 | 86 | 12 | 35 | 84 | 99 | 64 | 89 | 28 | 76 | 85 | 94 | 57 | 67 | 64 |
| 19 | 80 | 22 | 31 | 103 | 83 | 67 | 87 | 17 | 95 | 86 | 78 | 58 | 70 | 65 |
| 20 | 78 | 30 | 35 | 79 | 86 | 73 | 103 | 40 | 94 | 85 | 66 | 73 | 79 | 73 |
| 21 | 72 | 27 | 36 | 86 | 82 | 75 | 89 | 39 | 105 | 76 | 89 | 69 | 81 | 76 |
| 22 | 63 | 30 | 55 | 84 | 85 | 76 | 82 | 40 | 81 | 68 | 76 | 72 | 54 | 55 |
| 23 | 76 | 26 | 38 | 80 | 90 | 54 | 88 | 37 | 94 | 91 | 62 | 74 | 95 | 71 |
| 24 | 75 | 34 | 27 | 82 | 82 | 68 | 86 | 23 | 90 | 70 | 59 | 59 | 68 | 67 |
| 25 | 62 | 17 | 37 | 61 | 70 | 55 | 95 | 28 | 89 | 77 | 76 | 67 | 61 | 83 |
| 26 | 76 | 28 | 22 | 83 | 83 | 70 | 80 | 33 | 82 | 77 | 79 | 77 | 63 | 80 |
| 27 | 67 | 37 | 36 | 96 | 100 | 84 | 83 | 15 | 92 | 96 | 58 | 71 | 71 | 87 |
| 28 | 72 | 24 | 24 | 93 | 93 | 80 | 93 | 44 | 92 | 89 | 90 | 93 | 75 | 62 |
| 29 | 64 | 29 | 13 | 96 | 90 | 73 | 96 | 24 | 89 | 81 | 73 | 65 | 78 | 82 |
| 30 | 86 | 28 | 37 | 106 | 100 | 89 | 86 | 19 | 78 | 91 | 85 | 76 | 69 | 54 |
| 31 | 71 | 27 | 37 | 91 | 94 | 74 | 100 | 27 | 86 | 87 | 67 | 77 | 80 | 68 |
| 32 | 82 | 21 | 40 | 91 | 89 | 75 | 77 | 42 | 76 | 74 | 65 | 71 | 62 | 73 |
| 33 | 74 | 19 | 38 | 83 | 80 | 61 | 87 | 36 | 77 | 81 | 78 | 76 | 71 | 54 |
| 34 | 62 | 18 | 35 | 88 | 103 | 82 | 97 | 27 | 92 | 92 | 87 | 75 | 64 | 71 |
| 35 | 70 | 31 | 32 | 81 | 78 | 70 | 95 | 16 | 90 | 72 | 75 | 67 | 84 | 84 |
| 36 | 54 | 27 | 28 | 67 | 69 | 61 | 72 | 27 | 83 | 69 | 57 | 64 | 74 | 79 |
| 37 | 73 | 22 | 30 | 78 | 83 | 73 | 93 | 31 | 76 | 64 | 82 | 66 | 66 | 65 |
| 38 | 70 | 47 | 29 | 89 | 99 | 68 | 63 | 9 | 64 | 81 | 49 | 65 | 58 | 60 |
| 39 | 68 | 31 | 19 | 78 | 78 | 68 | 93 | 37 | 94 | 82 | 64 | 71 | 71 | 78 |
| 40 | 70 | 26 | 34 | 83 | 90 | 57 | 82 | 26 | 94 | 76 | 83 | 62 | 98 | 83 |
| 41 | 80 | 38 | 45 | 78 | 69 | 67 | 62 | 34 | 54 | 72 | 70 | 67 | 44 | 46 |
| 42 | 68 | 37 | 26 | 95 | 83 | 69 | 70 | 39 | 82 | 83 | 67 | 72 | 74 | 71 |
| 43 | 65 | 26 | 36 | 85 | 80 | 71 | 76 | 32 | 69 | 78 | 59 | 62 | 69 | 55 |
| 44 | 60 | 28 | 19 | 74 | 90 | 66 | 86 | 19 | 88 | 93 | 88 | 75 | 75 | 68 |
| 45 | 57 | 20 | 43 | 82 | 101 | 82 | 73 | 24 | 69 | 104 | 70 | 71 | 68 | 57 |
| 46 | 63 | 23 | 32 | 85 | 92 | 56 | 84 | 9 | 82 | 90 | 58 | 72 | 58 | 84 |
| 47 | 48 | 25 | 28 | 82 | 86 | 67 | 86 | 20 | 83 | 98 | 75 | 75 | 66 | 71 |
| 48 | 94 | 34 | 37 | 103 | 116 | 92 | 92 | 15 | 87 | 94 | 69 | 81 | 63 | 67 |
| 49 | 79 | 35 | 30 | 95 | 79 | 59 | 78 | 27 | 84 | 91 | 59 | 76 | 63 | 65 |
| 50 | 60 | 30 | 21 | 86 | 81 | 75 | 100 | 23 | 86 | 77 | 64 | 65 | 72 | 78 |

DATA-SIMULATIE VI
 DRIE GROEPSFACTOREN EN EEN SUBFACTOR
 (ZIE SIMULATIE VI POPULATIE FACTORSTRUCTUUR)
 VOOR 14 OPPERHATEN BIJ 100 ELEMENTEN

VERVOLG

| | | | | | | | | | | | | | | |
|----|----|----|----|-----|-----|----|-----|----|-----|-----|----|----|----|-----|
| 01 | 05 | 29 | 37 | 42 | 87 | 74 | 81 | 24 | 84 | 80 | 74 | 62 | 38 | 82 |
| 02 | 92 | 31 | 45 | 88 | 88 | 71 | 88 | 15 | 89 | 87 | 63 | 05 | 67 | 76 |
| 03 | 05 | 15 | 55 | 92 | 71 | 67 | 56 | 32 | 90 | 74 | 83 | 90 | 79 | 64 |
| 04 | 04 | 25 | 53 | 78 | 92 | 54 | 80 | 23 | 92 | 84 | 66 | 78 | 73 | 75 |
| 05 | 55 | 16 | 26 | 91 | 86 | 57 | 92 | 24 | 77 | 75 | 59 | 72 | 58 | 65 |
| 06 | 66 | 27 | 32 | 70 | 79 | 69 | 78 | 10 | 74 | 78 | 57 | 67 | 57 | 71 |
| 07 | 85 | 50 | 57 | 107 | 111 | 75 | 81 | 25 | 96 | 97 | 88 | 71 | 67 | 73 |
| 08 | 33 | 50 | 44 | 97 | 95 | 74 | 79 | 19 | 75 | 82 | 64 | 62 | 71 | 68 |
| 09 | 70 | 27 | 38 | 82 | 87 | 67 | 86 | 27 | 95 | 80 | 64 | 72 | 74 | 56 |
| 10 | 82 | 26 | 45 | 100 | 101 | 86 | 103 | 21 | 87 | 90 | 80 | 79 | 79 | 85 |
| 11 | 55 | 11 | 22 | 70 | 77 | 57 | 101 | 15 | 99 | 107 | 78 | 86 | 82 | 77 |
| 12 | 70 | 16 | 23 | 78 | 77 | 54 | 86 | 31 | 90 | 85 | 66 | 67 | 82 | 60 |
| 13 | 43 | 19 | 55 | 74 | 60 | 59 | 64 | 35 | 66 | 71 | 69 | 02 | 54 | 52 |
| 14 | 70 | 33 | 29 | 85 | 79 | 77 | 99 | 22 | 90 | 84 | 62 | 67 | 79 | 80 |
| 15 | 80 | 40 | 47 | 90 | 100 | 69 | 85 | 25 | 94 | 77 | 75 | 06 | 65 | 74 |
| 16 | 61 | 36 | 25 | 75 | 74 | 69 | 100 | 15 | 87 | 61 | 63 | 50 | 87 | 75 |
| 17 | 57 | 30 | 25 | 87 | 75 | 58 | 71 | 15 | 77 | 84 | 66 | 69 | 81 | 60 |
| 18 | 05 | 44 | 30 | 73 | 86 | 55 | 87 | 17 | 84 | 86 | 64 | 52 | 73 | 74 |
| 19 | 57 | 18 | 51 | 91 | 93 | 61 | 72 | 13 | 76 | 84 | 80 | 77 | 51 | 58 |
| 20 | 74 | 29 | 44 | 71 | 81 | 75 | 83 | 20 | 91 | 82 | 73 | 72 | 51 | 68 |
| 21 | 71 | 31 | 26 | 85 | 86 | 72 | 83 | 20 | 97 | 81 | 66 | 82 | 76 | 79 |
| 22 | 64 | 20 | 31 | 96 | 71 | 65 | 81 | 21 | 85 | 81 | 74 | 65 | 55 | 57 |
| 23 | 65 | 9 | 30 | 02 | 51 | 48 | 87 | 51 | 99 | 76 | 54 | 65 | 75 | 74 |
| 24 | 61 | 20 | 28 | 92 | 81 | 71 | 93 | 30 | 92 | 81 | 61 | 68 | 70 | 82 |
| 25 | 61 | 35 | 41 | 84 | 90 | 92 | 88 | 20 | 80 | 75 | 69 | 80 | 67 | 65 |
| 26 | 79 | 36 | 35 | 85 | 100 | 77 | 73 | 4 | 90 | 83 | 48 | 68 | 80 | 67 |
| 27 | 73 | 27 | 21 | 72 | 76 | 76 | 69 | 35 | 84 | 64 | 74 | 61 | 67 | 61 |
| 28 | 54 | 11 | 34 | 73 | 72 | 55 | 92 | 29 | 75 | 77 | 82 | 75 | 58 | 52 |
| 29 | 70 | 35 | 32 | 85 | 93 | 65 | 89 | 14 | 97 | 85 | 62 | 59 | 68 | 80 |
| 30 | 72 | 25 | 21 | 73 | 74 | 77 | 91 | 32 | 82 | 83 | 81 | 72 | 79 | 75 |
| 31 | 65 | 44 | 51 | 92 | 89 | 80 | 63 | 26 | 60 | 85 | 59 | 61 | 70 | 56 |
| 32 | 59 | 17 | 36 | 72 | 79 | 71 | 96 | 38 | 76 | 85 | 97 | 82 | 58 | 72 |
| 33 | 62 | 42 | 26 | 82 | 83 | 75 | 85 | 16 | 90 | 80 | 54 | 68 | 75 | 74 |
| 34 | 69 | 28 | 37 | 80 | 89 | 71 | 85 | 20 | 76 | 85 | 64 | 62 | 62 | 63 |
| 35 | 60 | 20 | 15 | 89 | 77 | 64 | 87 | 28 | 82 | 82 | 56 | 94 | 83 | 77 |
| 36 | 55 | 21 | 22 | 90 | 68 | 55 | 95 | 31 | 93 | 95 | 54 | 79 | 74 | 100 |
| 37 | 71 | 16 | 52 | 60 | 73 | 76 | 80 | 27 | 70 | 34 | 65 | 61 | 62 | 59 |
| 38 | 61 | 29 | 42 | 90 | 64 | 59 | 59 | 19 | 86 | 71 | 73 | 70 | 63 | 71 |
| 39 | 68 | 15 | 34 | 100 | 91 | 63 | 93 | 27 | 84 | 104 | 88 | 77 | 60 | 68 |
| 40 | 70 | 11 | 32 | 71 | 75 | 46 | 93 | 18 | 91 | 90 | 63 | 78 | 61 | 72 |
| 41 | 84 | 30 | 37 | 88 | 83 | 71 | 81 | 15 | 80 | 96 | 70 | 71 | 66 | 76 |
| 42 | 83 | 31 | 36 | 92 | 89 | 78 | 106 | 25 | 102 | 94 | 63 | 78 | 84 | 95 |
| 43 | 64 | 27 | 37 | 88 | 95 | 81 | 84 | 32 | 81 | 75 | 61 | 67 | 65 | 68 |
| 44 | 75 | 23 | 51 | 90 | 95 | 75 | 89 | 30 | 72 | 81 | 71 | 75 | 58 | 61 |
| 45 | 79 | 31 | 21 | 88 | 93 | 71 | 89 | 28 | 82 | 76 | 55 | 63 | 73 | 72 |
| 46 | 65 | 27 | 28 | 64 | 78 | 62 | 90 | 18 | 100 | 89 | 84 | 76 | 70 | 83 |
| 47 | 60 | 16 | 24 | 70 | 69 | 60 | 95 | 33 | 94 | 78 | 76 | 67 | 59 | 71 |
| 48 | 82 | 23 | 30 | 83 | 82 | 67 | 89 | 31 | 87 | 91 | 73 | 78 | 77 | 73 |
| 49 | 73 | 35 | 42 | 81 | 76 | 74 | 80 | 25 | 84 | 70 | 63 | 71 | 62 | 76 |
| 50 | 77 | 32 | 31 | 94 | 92 | 72 | 90 | 39 | 100 | 78 | 52 | 76 | 71 | 68 |

DATASIMULATIE VII
 DRIE GROEPSFAKTOREN EN EEN SUBFAKTOR
 (ZIE SIMULATIE VI POPULATIE FAKTORSTRUKTUUR)
 VOOR 14 OPERATOREN BIJ 100 ELEMENTEN

| | | | | | | | | | | | | | | |
|----|----|----|----|-----|-----|----|-----|----|-----|-----|----|----|----|----|
| 1 | 58 | 21 | 45 | 71 | 87 | 58 | 87 | 38 | 86 | 20 | 86 | 72 | 57 | 80 |
| 2 | 59 | 18 | 26 | 79 | 74 | 65 | 88 | 33 | 78 | 87 | 67 | 70 | 59 | 65 |
| 3 | 73 | 35 | 35 | 97 | 95 | 76 | 88 | 24 | 64 | 90 | 68 | 77 | 68 | 66 |
| 4 | 90 | 43 | 43 | 90 | 98 | 81 | 97 | 24 | 79 | 80 | 63 | 63 | 68 | 70 |
| 5 | 85 | 41 | 30 | 92 | 94 | 71 | 77 | 27 | 75 | 76 | 63 | 67 | 71 | 72 |
| 6 | 84 | 21 | 22 | 91 | 76 | 79 | 93 | 12 | 98 | 90 | 70 | 63 | 76 | 82 |
| 7 | 62 | 13 | 22 | 70 | 81 | 54 | 79 | 19 | 64 | 78 | 60 | 59 | 74 | 70 |
| 8 | 64 | 19 | 57 | 78 | 78 | 73 | 84 | 18 | 82 | 72 | 67 | 61 | 61 | 64 |
| 9 | 71 | 35 | 36 | 110 | 102 | 77 | 85 | 13 | 90 | 86 | 60 | 66 | 74 | 89 |
| 10 | 56 | 55 | 52 | 74 | 71 | 75 | 78 | 38 | 68 | 64 | 66 | 62 | 53 | 66 |
| 11 | 80 | 19 | 31 | 68 | 76 | 69 | 82 | 44 | 90 | 77 | 76 | 74 | 60 | 66 |
| 12 | 65 | 9 | 37 | 78 | 73 | 56 | 100 | 34 | 87 | 84 | 85 | 65 | 68 | 78 |
| 13 | 69 | 4 | 43 | 90 | 87 | 56 | 88 | 27 | 83 | 105 | 73 | 92 | 62 | 62 |
| 14 | 63 | 40 | 40 | 103 | 88 | 68 | 65 | 9 | 62 | 72 | 57 | 54 | 65 | 57 |
| 15 | 77 | 25 | 30 | 80 | 85 | 80 | 92 | 27 | 67 | 88 | 74 | 84 | 71 | 79 |
| 16 | 73 | 17 | 27 | 77 | 80 | 61 | 80 | 25 | 80 | 96 | 71 | 72 | 67 | 56 |
| 17 | 73 | 22 | 45 | 78 | 72 | 73 | 76 | 41 | 84 | 73 | 79 | 79 | 66 | 73 |
| 18 | 76 | 10 | 29 | 90 | 88 | 67 | 102 | 28 | 109 | 95 | 62 | 73 | 79 | 80 |
| 19 | 73 | 28 | 40 | 67 | 75 | 75 | 94 | 39 | 88 | 60 | 78 | 81 | 83 | 80 |
| 20 | 66 | 25 | 26 | 95 | 95 | 77 | 108 | 33 | 93 | 93 | 69 | 85 | 90 | 94 |
| 21 | 82 | 32 | 25 | 82 | 84 | 74 | 106 | 36 | 92 | 81 | 60 | 78 | 70 | 68 |
| 22 | 77 | 29 | 26 | 83 | 84 | 69 | 96 | 36 | 71 | 76 | 68 | 70 | 72 | 61 |
| 23 | 71 | 35 | 44 | 78 | 77 | 59 | 73 | 25 | 67 | 71 | 64 | 69 | 67 | 69 |
| 24 | 62 | 12 | 34 | 77 | 71 | 52 | 88 | 41 | 94 | 82 | 80 | 74 | 75 | 72 |
| 25 | 93 | 24 | 25 | 91 | 103 | 81 | 106 | 22 | 96 | 102 | 84 | 91 | 76 | 91 |
| 26 | 67 | 40 | 28 | 84 | 91 | 82 | 95 | 32 | 83 | 83 | 58 | 64 | 81 | 86 |
| 27 | 85 | 39 | 17 | 113 | 108 | 79 | 100 | 4 | 92 | 95 | 61 | 62 | 86 | 83 |
| 28 | 78 | 41 | 41 | 111 | 107 | 75 | 81 | 32 | 73 | 93 | 43 | 65 | 61 | 61 |
| 29 | 75 | 35 | 42 | 82 | 89 | 72 | 82 | 35 | 80 | 85 | 77 | 79 | 53 | 65 |
| 30 | 74 | 27 | 26 | 93 | 87 | 71 | 94 | 44 | 91 | 84 | 82 | 80 | 85 | 74 |
| 31 | 67 | 37 | 36 | 78 | 78 | 64 | 77 | 17 | 70 | 81 | 68 | 68 | 77 | 61 |
| 32 | 74 | 45 | 46 | 90 | 80 | 79 | 78 | 34 | 87 | 76 | 72 | 65 | 63 | 81 |
| 33 | 68 | 32 | 45 | 91 | 84 | 82 | 97 | 23 | 96 | 97 | 68 | 73 | 69 | 63 |
| 34 | 83 | 35 | 49 | 98 | 95 | 72 | 76 | 28 | 76 | 98 | 74 | 73 | 60 | 67 |
| 35 | 81 | 20 | 22 | 76 | 96 | 71 | 68 | 31 | 81 | 85 | 57 | 65 | 71 | 76 |
| 36 | 81 | 29 | 33 | 76 | 77 | 67 | 68 | 29 | 75 | 72 | 56 | 54 | 59 | 55 |
| 37 | 57 | 23 | 24 | 68 | 78 | 63 | 90 | 11 | 96 | 82 | 60 | 60 | 78 | 82 |
| 38 | 77 | 28 | 32 | 85 | 89 | 71 | 87 | 16 | 83 | 98 | 52 | 71 | 83 | 80 |
| 39 | 57 | 18 | 16 | 101 | 95 | 49 | 107 | 71 | 105 | 91 | 62 | 72 | 87 | 82 |
| 40 | 73 | 19 | 19 | 74 | 89 | 66 | 99 | 20 | 94 | 97 | 75 | 87 | 70 | 79 |
| 41 | 59 | 17 | 27 | 91 | 79 | 60 | 73 | 15 | 86 | 84 | 64 | 60 | 71 | 81 |
| 42 | 77 | 24 | 38 | 82 | 85 | 73 | 99 | 39 | 93 | 68 | 65 | 79 | 85 | 84 |
| 43 | 69 | 29 | 53 | 73 | 74 | 72 | 75 | 30 | 69 | 79 | 62 | 72 | 49 | 62 |
| 44 | 61 | 36 | 24 | 91 | 95 | 70 | 79 | 29 | 87 | 82 | 74 | 67 | 60 | 67 |
| 45 | 61 | 29 | 38 | 93 | 91 | 83 | 88 | 30 | 89 | 80 | 57 | 68 | 86 | 69 |
| 46 | 70 | 15 | 25 | 76 | 82 | 74 | 95 | 38 | 88 | 86 | 73 | 69 | 70 | 72 |
| 47 | 92 | 24 | 32 | 83 | 89 | 66 | 89 | 21 | 69 | 81 | 71 | 55 | 64 | 58 |
| 48 | 79 | 30 | 19 | 81 | 87 | 63 | 83 | 24 | 82 | 80 | 71 | 69 | 63 | 77 |
| 49 | 40 | 12 | 15 | 87 | 81 | 65 | 85 | 27 | 67 | 90 | 63 | 73 | 58 | 69 |
| 50 | 73 | 13 | 32 | 75 | 80 | 59 | 81 | 31 | 97 | 86 | 73 | 63 | 54 | 90 |

DATASIMULATIE VII
 DEEL GROEPSFACTOREN EN EEN SUBFAKTOR
 (ZIE SIMULATIE VI POPULATIE FAKTORSTRUKTUUR)
 VOOR 14 OPERATOREN EIJ 100 ELEMENTEN

VERVOLG

| | | | | | | | | | | | | | | |
|-----|----|----|----|-----|-----|----|-----|----|-----|-----|----|----|----|----|
| 51 | 72 | 24 | 32 | 95 | 83 | 65 | 83 | 34 | 75 | 80 | 69 | 61 | 60 | 65 |
| 52 | 71 | 21 | 13 | 82 | 86 | 72 | 98 | 28 | 79 | 78 | 65 | 72 | 54 | 58 |
| 53 | 81 | 38 | 26 | 73 | 88 | 75 | 83 | 14 | 82 | 79 | 66 | 53 | 64 | 62 |
| 54 | 79 | 27 | 49 | 109 | 102 | 56 | 71 | 25 | 91 | 88 | 65 | 66 | 63 | 81 |
| 55 | 82 | 29 | 33 | 91 | 101 | 82 | 84 | 21 | 84 | 83 | 63 | 70 | 64 | 69 |
| 56 | 71 | 34 | 36 | 91 | 87 | 61 | 89 | 21 | 45 | 92 | 64 | 76 | 76 | 90 |
| 57 | 59 | 24 | 30 | 92 | 83 | 73 | 88 | 18 | 97 | 91 | 60 | 66 | 71 | 71 |
| 58 | 54 | 17 | 21 | 61 | 69 | 48 | 80 | 30 | 87 | 75 | 79 | 71 | 71 | 76 |
| 59 | 50 | 13 | 28 | 67 | 82 | 60 | 79 | 30 | 85 | 95 | 61 | 76 | 60 | 73 |
| 60 | 77 | 16 | 41 | 84 | 85 | 77 | 62 | 46 | 10 | 82 | 83 | 84 | 82 | 68 |
| 61 | 65 | 35 | 34 | 91 | 93 | 71 | 78 | 10 | 76 | 95 | 68 | 66 | 56 | 64 |
| 62 | 71 | 29 | 12 | 73 | 78 | 60 | 99 | 35 | 104 | 88 | 76 | 78 | 80 | 78 |
| 63 | 80 | 47 | 55 | 97 | 96 | 72 | 89 | 18 | 80 | 96 | 60 | 64 | 72 | 70 |
| 64 | 50 | 27 | 28 | 81 | 79 | 60 | 82 | 26 | 77 | 81 | 56 | 61 | 70 | 85 |
| 65 | 73 | 38 | 29 | 62 | 83 | 67 | 74 | 8 | 78 | 69 | 65 | 61 | 70 | 65 |
| 66 | 72 | 39 | 24 | 83 | 92 | 75 | 78 | 15 | 76 | 34 | 63 | 68 | 68 | 69 |
| 67 | 55 | 25 | 17 | 80 | 75 | 59 | 87 | 27 | 90 | 89 | 53 | 65 | 82 | 81 |
| 68 | 73 | 25 | 49 | 88 | 88 | 38 | 80 | 18 | 65 | 79 | 65 | 63 | 65 | 72 |
| 69 | 61 | 26 | 22 | 82 | 91 | 67 | 77 | 8 | 70 | 84 | 56 | 69 | 63 | 67 |
| 70 | 83 | 36 | 41 | 84 | 98 | 78 | 75 | 4 | 92 | 85 | 69 | 57 | 58 | 56 |
| 71 | 61 | 32 | 25 | 94 | 93 | 56 | 91 | 26 | 79 | 31 | 76 | 53 | 82 | 76 |
| 72 | 77 | 21 | 36 | 96 | 75 | 55 | 97 | 33 | 102 | 83 | 75 | 81 | 76 | 83 |
| 73 | 65 | 24 | 21 | 91 | 98 | 31 | 98 | 35 | 65 | 46 | 85 | 72 | 84 | 64 |
| 74 | 48 | 18 | 28 | 96 | 79 | 50 | 98 | 8 | 80 | 96 | 64 | 58 | 81 | 81 |
| 75 | 66 | 36 | 27 | 80 | 73 | 52 | 68 | 11 | 74 | 78 | 58 | 46 | 80 | 66 |
| 76 | 61 | 33 | 34 | 92 | 79 | 78 | 92 | 31 | 61 | 73 | 75 | 73 | 88 | 69 |
| 77 | 77 | 26 | 25 | 92 | 101 | 64 | 86 | 21 | 107 | 90 | 41 | 55 | 83 | 83 |
| 78 | 52 | 11 | 28 | 64 | 75 | 59 | 86 | 37 | 88 | 86 | 75 | 77 | 63 | 75 |
| 79 | 88 | 29 | 43 | 99 | 98 | 65 | 93 | 20 | 81 | 105 | 88 | 82 | 56 | 69 |
| 80 | 66 | 5 | 36 | 85 | 86 | 73 | 83 | 26 | 83 | 101 | 83 | 81 | 64 | 64 |
| 81 | 84 | 25 | 26 | 87 | 85 | 72 | 95 | 38 | 92 | 88 | 78 | 75 | 88 | 90 |
| 82 | 60 | 16 | 34 | 66 | 73 | 55 | 78 | 32 | 77 | 86 | 86 | 86 | 53 | 46 |
| 83 | 68 | 19 | 38 | 80 | 78 | 49 | 83 | 23 | 81 | 98 | 63 | 71 | 57 | 67 |
| 84 | 71 | 32 | 53 | 103 | 98 | 75 | 66 | 16 | 81 | 91 | 69 | 70 | 74 | 72 |
| 85 | 75 | 18 | 26 | 78 | 85 | 75 | 102 | 39 | 93 | 87 | 74 | 82 | 68 | 80 |
| 86 | 73 | 3 | 43 | 77 | 76 | 74 | 94 | 26 | 95 | 86 | 79 | 79 | 62 | 72 |
| 87 | 69 | 23 | 28 | 87 | 87 | 63 | 90 | 12 | 86 | 98 | 56 | 72 | 73 | 82 |
| 88 | 79 | 37 | 47 | 103 | 90 | 75 | 78 | 24 | 71 | 95 | 74 | 65 | 65 | 53 |
| 89 | 63 | 33 | 44 | 83 | 80 | 81 | 97 | 30 | 71 | 38 | 55 | 60 | 68 | 67 |
| 90 | 72 | 31 | 39 | 75 | 94 | 71 | 82 | 41 | 65 | 78 | 76 | 73 | 68 | 51 |
| 91 | 74 | 34 | 41 | 87 | 75 | 74 | 79 | 28 | 72 | 78 | 67 | 60 | 62 | 65 |
| 92 | 81 | 26 | 50 | 81 | 93 | 78 | 83 | 34 | 100 | 93 | 97 | 96 | 62 | 56 |
| 93 | 82 | 26 | 37 | 88 | 84 | 75 | 84 | 41 | 87 | 51 | 81 | 81 | 72 | 60 |
| 94 | 55 | 41 | 50 | 82 | 76 | 62 | 72 | 23 | 79 | 76 | 62 | 59 | 74 | 67 |
| 95 | 79 | 42 | 43 | 82 | 94 | 82 | 73 | 14 | 71 | 88 | 71 | 60 | 69 | 54 |
| 96 | 56 | 29 | 49 | 70 | 86 | 82 | 80 | 32 | 78 | 85 | 79 | 69 | 68 | 59 |
| 97 | 75 | 28 | 19 | 78 | 69 | 69 | 83 | 30 | 84 | 90 | 62 | 63 | 74 | 79 |
| 98 | 73 | 9 | 34 | 82 | 94 | 69 | 99 | 42 | 81 | 117 | 74 | 92 | 66 | 52 |
| 99 | 70 | 35 | 43 | 87 | 86 | 67 | 79 | 24 | 74 | 82 | 63 | 63 | 46 | 64 |
| 100 | 62 | 39 | 45 | 75 | 86 | 85 | 90 | 29 | 77 | 86 | 73 | 80 | 65 | 52 |

DATASIMULATIE VIII
 DRIE GROEPSKAKTOREN EN DRIE SUBFAKTOREN
 (ZIE TABEL 6.3.2.1.)
 VOOR 14 OPERATOREN BIJ 100 ELEMENTEN

| | | | | | | | | | |
|----|--------|-------|----------|-------------------|----------------|----------------|----------------|-------------|----------|
| 1 | 82 | 91 | 64100 | 59 | 93110 | 67104108 | 84103 | 84 | 93 |
| 2 | 91112 | 55107 | 71 | 88102 | 44103 | 98107 | 83 | 94 | 86 |
| 3 | 104113 | 79109 | 55 | 95 | 96 | 45110100 | 99 | 87 | 75 |
| 4 | 94126 | 58124 | 78 | 88108 | 53110103 | 95 | 84 | 85 | 90 |
| 5 | 99100 | 61111 | 65 | 97106 | 68108 | 89108 | 97 | 82 | 91 |
| 6 | 92 | 97 | 59101 | 70 | 86112 | 75113117105104 | 80 | 78 | |
| 7 | 104102 | 68105 | 57 | 82113 | 70111 | 99112 | 94 | 77 | 85 |
| 8 | 97116 | 68113 | 47 | 90103 | 45103108 | 84 | 79102102 | | |
| 9 | 96 | 89 | 52104 | 45 | 68105 | 66 | 95107101 | 97103 | 87 |
| 10 | 123118 | 60119 | 73103123 | 64125123103108111 | 96 | | | | |
| 11 | 77107 | 52114 | 65 | 93106 | 70111114111 | 93 | 91 | 96 | |
| 12 | 89120 | 58121 | 59 | 87123 | 71120119114111 | 98 | 96 | | |
| 13 | 91117 | 59124 | 84 | 97123 | 56102113106105 | 92 | 95 | | |
| 14 | 91 | 90 | 58 | 98 | 60 | 78110 | 65106114107102 | 84 | 93 |
| 15 | 102107 | 59 | 99 | 67 | 90 | 93 | 57111 | 90 | 82 |
| 16 | 95106 | 49102 | 53 | 76107 | 42 | 98 | 90 | 89 | 78 |
| 17 | 97116 | 55111 | 70 | 96106 | 55109123 | 99105 | 86 | 96 | |
| 18 | 97100 | 54 | 93 | 52 | 86108 | 43 | 92 | 96106 | 85 |
| 19 | 65112 | 62110 | 62 | 74 | 94 | 53 | 81105 | 94 | 89 |
| 20 | 82 | 93 | 50 | 81 | 63 | 70 | 92 | 73110114103 | 95 |
| 21 | 98120 | 63111 | 76104119 | 66116118107109 | 88102 | | | | |
| 22 | 89109 | 40120 | 59 | 84113 | 50 | 91100 | 92 | 94 | 96102 |
| 23 | 101120 | 77110 | 72 | 90102 | 70106118 | 87 | 94 | 78 | 88 |
| 24 | 85 | 99 | 58106 | 55 | 89113 | 56 | 96100 | 87 | 84 |
| 25 | 97105 | 53108 | 56 | 80 | 97 | 45 | 96 | 93 | 78 |
| 26 | 98100 | 60114 | 56 | 88110 | 68113109103100 | 90 | 87 | | |
| 27 | 98121 | 68118 | 51 | 83119 | 43112 | 80 | 75 | 80100104 | |
| 28 | 116119 | 56102 | 68102114 | 66125100100 | 97101 | 90 | | | |
| 29 | 107108 | 78107 | 43 | 90120 | 73 | 98100 | 99 | 96103 | 71 |
| 30 | 100102 | 62121 | 62104106 | 52 | 94112 | 99 | 89 | 90101 | |
| 31 | 97102 | 48115 | 59101116 | 70113 | 95103 | 88111115 | | | |
| 32 | 97107 | 69111 | 62 | 94115 | 70108110101 | 95 | 82 | 86 | |
| 33 | 102128 | 75131 | 72110 | 92 | 52116104103 | 80 | 79 | 92 | |
| 34 | 89100 | 54118 | 71 | 85102 | 56 | 98119 | 90 | 97 | 90 |
| 35 | 104121 | 53 | 94 | 63 | 96115 | 54106101 | 87 | 98 | 87100 |
| 36 | 98120 | 81115 | 82112103 | 56112101103 | 91 | 78 | 67 | | |
| 37 | 97 | 98 | 49103 | 50 | 90109 | 67101 | 95 | 84 | 93101117 |
| 38 | 91122 | 62110 | 67100108 | 49111114110 | 94 | 87 | 96 | | |
| 39 | 109112 | 51115 | 68100111 | 51110117102108 | 91104 | | | | |
| 40 | 103117 | 67132 | 62 | 92104 | 52108106 | 89 | 85 | 88 | 88 |
| 41 | 84100 | 70109 | 51 | 96 | 99 | 73 | 93 | 97 | 87 |
| 42 | 99112 | 66124 | 77101109 | 61121106108103 | 80 | 80 | | | |
| 43 | 95118 | 75115 | 47 | 76 | 99 | 67102116 | 80 | 94 | 75 |
| 44 | 103115 | 79107 | 79103 | 92 | 77107 | 98 | 98 | 91 | 81 |
| 45 | 107110 | 51113 | 67 | 83 | 95 | 48111 | 92 | 93 | 83 |
| 46 | 120106 | 59112 | 71121 | 94 | 58116 | 94 | 84 | 83 | 95 |
| 47 | 98104 | 34100 | 51 | 85120 | 64 | 96122114107 | 92 | 92 | |
| 48 | 96102 | 44107 | 54 | 90114 | 67113102104106 | 97 | 99 | | |
| 49 | 86101 | 75108 | 65 | 81 | 94 | 52113 | 97 | 90 | 91 |
| 50 | 105102 | 72101 | 57 | 90110 | 71111111111103 | 94 | 98 | | |

DATA-SIMULATIE VIII
 DRIE GROEPSKAKTOREN EN DRIE SUBFAKTOREN
 (ZIE TABEL 6.3.2.1.)
 VOOR 14 OPERATOREN BIJ 100 ELEMENTEN

VERVOLG

| | | | | | | | | | | | |
|-----|--------|-------|----------|----------------|----------------|-------------|----------------------|----------|----------|----------|----|
| 51 | 88 | 89 | 71113 | 55 | 85 | 94 | 54 | 97103104 | 94 | 89 | 94 |
| 52 | 100109 | 36 | 99 | 59 | 85129 | 48 | 98114 | 97108121 | 95 | | |
| 53 | 85113 | 46119 | 67 | 88113 | 47103125112106 | 83106 | | | | | |
| 54 | 111116 | 56107 | 73 | 95120 | 67112112 | 96100 | 81 | 95 | | | |
| 55 | 105117 | 56117 | 67102103 | 55107104 | 92 | 89 | 94 | 95 | | | |
| 56 | 101119 | 62123 | 56 | 99105 | 42 | 89107 | 79 | 82 | 96 | 96 | |
| 57 | 109116 | 56126 | 74104123 | 59126 | 98 | 93 | 90116108 | | | | |
| 58 | 108131 | 60127 | 64107108 | 38124115102 | 99107102 | | | | | | |
| 59 | 97108 | 47104 | 66 | 87113 | 82113124116119 | 93 | 87 | | | | |
| 60 | 93103 | 56114 | 69 | 84118 | 68104114115107 | 97106 | | | | | |
| 61 | 91108 | 55121 | 62 | 88106 | 66103108 | 97 | 87 | 89101 | | | |
| 62 | 92 | 95 | 26101 | 59 | 72120 | 60108110 | 97 | 95118113 | | | |
| 63 | 113115 | 47120 | 36 | 92 | 95 | 32106 | 94 | 79 | 78102 | 81 | |
| 64 | 79103 | 67101 | 59 | 96118 | 75110107100 | 99 | 90104 | | | | |
| 65 | 104121 | 74118 | 70 | 93111 | 52116 | 96 | 92 | 82 | 94101 | | |
| 66 | 107120 | 64125 | 62104112 | 59123104102 | 91105 | 97 | | | | | |
| 67 | 97115 | 66110 | 63 | 80102 | 57105120110 | 88 | 89 | 83 | | | |
| 68 | 91103 | 68105 | 72 | 93 | 89 | 60 | 97 | 94 | 93 | 80 | 89 |
| 69 | 92102 | 47 | 95 | 60 | 88112 | 64 | 96 | 83 | 88 | 82111107 | |
| 70 | 92108 | 62107 | 67 | 96124 | 68105100111 | 94 | 87 | 91 | | | |
| 71 | 91123 | 67119 | 75106119 | 59114105103 | 92111104 | | | | | | |
| 72 | 91 | 89 | 63120 | 64 | 95108 | 79 | 99100 | 93 | 95102 | 88 | |
| 73 | 93106 | 44 | 93 | 47 | 72109 | 49 | 94101106 | 93106 | 94 | | |
| 74 | 93 | 94 | 41 | 87 | 48 | 83 | 97 | 63 | 96101 | 88 | 84 |
| 75 | 90100 | 55106 | 70 | 84 | 95 | 49100 | 96 | 85 | 81 | 92 | 96 |
| 76 | 106104 | 50112 | 56 | 95101 | 65113101 | 89 | 97 | 89108 | | | |
| 77 | 90114 | 75117 | 68 | 86 | 90 | 51111 | 94 | 83 | 84 | 74 | 85 |
| 78 | 101122 | 68117 | 59 | 87 | 98 | 73110122 | 96100 | 89 | 80 | | |
| 79 | 108119 | 73110 | 65113 | 96 | 45116108 | 92 | 89 | 86 | 90 | | |
| 80 | 96117 | 73106 | 76101122 | 58117104112 | 98 | 97 | 99 | | | | |
| 81 | 85106 | 53109 | 65 | 82 | 99 | 67105 | 93 | 88 | 84 | 97109 | |
| 82 | 98118 | 79126 | 74110 | 96 | 59117 | 91 | 98 | 87 | 83 | 86 | |
| 83 | 92103 | 57106 | 76 | 95119 | 70108 | 91 | 84 | 80 | 95105 | | |
| 84 | 98105 | 55107 | 57 | 88115 | 61 | 99 | 98 | 96 | 86112113 | | |
| 85 | 90119 | 70119 | 56 | 86115 | 45104111101 | 88 | 89 | 86 | | | |
| 86 | 89109 | 66114 | 51 | 98114 | 63100103104104 | 90 | 86 | | | | |
| 87 | 95104 | 67115 | 69109114 | 71114 | 98103 | 92101102 | | | | | |
| 88 | 102120 | 65122 | 60107115 | 58109 | 95116 | 88 | 90 | 83 | | | |
| 89 | 95111 | 50114 | 60 | 84112 | 41106113 | 97 | 91 | 90 | 87 | | |
| 90 | 107 | 87 | 43106 | 58 | 94122 | 72114109 | 93102107107 | | | | |
| 91 | 88 | 91 | 57 | 87 | 50 | 82124 | 73111110105103105100 | | | | |
| 92 | 95 | 88 | 60103 | 54 | 90115 | 64106102119 | 93104 | 92 | | | |
| 93 | 105 | 99 | 62111 | 54 | 96123 | 60118104 | 95 | 93 | 88108 | | |
| 94 | 94111 | 60105 | 78105119 | 87116110119116 | 94 | 90 | | | | | |
| 95 | 100116 | 73113 | 61 | 94108 | 60 | 97102 | 92 | 87 | 92 | 86 | |
| 96 | 84116 | 42113 | 61 | 83 | 93 | 59107 | 99 | 96 | 94 | 92 | 95 |
| 97 | 95112 | 47108 | 61 | 87111 | 60102 | 83 | 87 | 84 | 93101 | | |
| 98 | 122122 | 71140 | 61114100 | 47114100104 | 84 | 99 | 90 | | | | |
| 99 | 95118 | 76116 | 73 | 99111 | 76117118110105 | 76 | 73 | | | | |
| 100 | 95122 | 69111 | 56 | 85 | 96 | 50103108 | 94 | 98 | 78 | 76 | |

BIJLAGE 11.2. DATA ZUSTERS (ZIE PARAGRAF 4.1.)

CODERING 1 : NEE-ANTWOORD
 CODERING 2 : GEEN MENING
 CODERING 3 : JA-ANTWOORD
 CODERING 5 OF 9 : GEEN-ANTWOORD

GROEP 1 : DE ANTWOORDEN OP 15 VRAGEN
 (ZIE TABEL 4.1.2.) VAN 120 ZUSTERS

| | | |
|-----------------|-----------------|------------------|
| 23213332221532 | 313311213331313 | 313113311332212 |
| 312113211333312 | 312122211233213 | 3121112221321222 |
| 313113311333313 | 312212212233313 | 313112311332212 |
| 133233111321222 | 312111331133313 | 312113311333212 |
| 311113311333313 | 322113213211233 | 313312311332212 |
| 313113211233313 | 322213212232313 | 312222212222222 |
| 112112113232222 | 332122213331213 | 312112212223322 |
| 323112312333322 | 313113311333312 | 312112111222212 |
| 313112333332212 | 313113311333313 | 322112311332212 |
| 312113311333313 | 313013312233313 | 322112211332212 |
| 313113311333312 | 313113311333323 | 313113311233212 |
| 222228922222222 | 312113311333313 | 313113311333312 |
| 313113221333213 | 313113311331312 | 312113311332212 |
| 322322222221312 | 313113311333313 | 131321123221122 |
| 313112311233313 | 313113311232313 | 348113311332212 |
| 303013311233313 | 312112212331213 | 322122211332222 |
| 313113311333213 | 313112311331323 | 313112311332212 |
| 333212211241313 | 313113211333312 | 312113311333212 |
| 313113311233313 | 312113311333312 | 311111113322212 |
| 313113311333312 | 313113311333313 | 313113311332212 |
| 313113311331213 | 313123321333313 | 312312311333313 |
| 318113113333312 | 312112213312222 | 312113311331313 |
| 311111111331313 | 313113311233313 | 312112311332212 |
| 313119811331322 | 312113312332313 | 313113311332212 |
| 312113212333212 | 313212311333233 | 313112211333312 |
| 313113311332212 | 313113311231312 | 312212233122221 |
| 313113311233313 | 313113311333312 | 111191133923111 |
| 112131113313213 | 111131113333313 | 311112212321312 |
| 111111113313132 | 312112212332321 | 313113311332212 |
| 312212212321313 | 313313211231313 | 311113311332212 |
| 21222221231312 | 223113211331223 | 312222212322212 |
| 311111113223113 | 312313211233213 | 313113311332212 |
| 31111111323113 | 313113222233312 | 222222222222223 |
| 313112311333313 | 313113311333312 | 313112322322212 |
| 313113311331313 | 313112211233312 | 303212222222222 |
| 123133311231311 | 313113311331312 | 313112311321313 |
| 322212212331323 | 313113211231313 | 313113311332212 |
| 313213212233133 | 123311211331112 | 312113311232212 |
| 222222211232212 | 312313312323322 | 322313123321222 |
| 213313311331312 | 313113311333313 | 312113311331312 |

GRUPP 2 : DE ANTWOORDEN OP 15 VRAGEN
(ZIE TABEL 4.1.2.) VAN 120 ZUSERS

| | | |
|-----------------|-----------------|-----------------|
| 313212312331313 | 313112311233313 | 113333311331312 |
| 313113311333313 | 318119819231313 | 313113211231313 |
| 313112311333313 | 313112211332313 | 313113311333313 |
| 313113311231313 | 318113312333323 | 312113311333313 |
| 323213321221922 | 313113311333313 | 223122211211312 |
| 313113311232313 | 311311111311113 | 222222222222222 |
| 323223212331213 | 212111212332323 | 333313311331313 |
| 311113111323313 | 318119811833313 | 313113311331313 |
| 311113311331313 | 319118911311313 | 322922222232323 |
| 313113311331313 | 223221121932132 | 313113311332313 |
| 313122222232222 | 132211212331122 | 311113311333313 |
| 332312213331233 | 313113311333313 | 312113212223333 |
| 313118811333313 | 223113311231313 | 313113311331313 |
| 222222222222222 | 113111311333113 | 213213312231112 |
| 313113311333313 | 313112211233313 | 313113211323313 |
| 312113211233313 | 313113311333313 | 333313311333313 |
| 313112311333313 | 313113211232213 | 313113311333313 |
| 313113311233313 | 313223312331313 | 313113311332313 |
| 313119911383313 | 313113312323313 | 313112313332223 |
| 333112212331333 | 222322213231122 | 323323211321213 |
| 322123913333223 | 132339913131213 | 223222222232222 |
| 22332331331312 | 111831113331111 | 313111311333313 |
| 313123312333313 | 322112311331212 | 319113911833313 |
| 313112222232323 | 322222223332313 | 222122212212222 |
| 22222222321223 | 313113211232313 | 33212223323312 |
| 312812213332313 | 312118911333313 | 313119891331313 |
| 312311311231113 | 313113311331313 | 312113211333313 |
| 21221223232123 | 222222222222222 | 313112211331313 |
| 311118913331113 | 213113311331313 | 313113313333313 |
| 323213311331313 | 312112212333313 | 322232222227223 |
| 312113312332313 | 993319311331311 | 333113311231313 |
| 312113311232313 | 313913911381393 | 312313311333313 |
| 312113311393813 | 323323311333323 | 223112312232323 |
| 312113212313313 | 392112212333313 | 113111311333113 |
| 323313211232313 | 313119311831313 | 312113113333213 |
| 313119911333313 | 313113311222313 | 313113311333313 |
| 313113211333313 | 313113311333313 | 313313313331313 |
| 311113311331313 | 313113311332313 | 313113311333313 |
| 332113311231313 | 313119911331313 | 313113311333313 |
| 313113311333333 | 223223222932223 | 313113311333313 |

GRUPE 3 : DE ANTWOORDEN OP 15 VRAGEN
(ZIE TABEL 4.1.2.) VAN 100 ZUSTERS

| | |
|-----------------|-----------------|
| 513113311333313 | 338311311331333 |
| 312112321231222 | 313113311333313 |
| 319113819333313 | 332332212221333 |
| 313113311331313 | 333113311331313 |
| 222222222222222 | 313113311331313 |
| 312123211331313 | 311119911311313 |
| 313113211331313 | 312113311333313 |
| 313134898333313 | 233322323231213 |
| 312113311332313 | 313113311333313 |
| 133331333331312 | 333313211331313 |
| 313113313233213 | 312118911231131 |
| 323213311332313 | 313113311332213 |
| 132111113231313 | 333222212331212 |
| 312118918833313 | 313113311333313 |
| 313113312333313 | 212111133332121 |
| 312111211333313 | 312118913311113 |
| 332332232333212 | 313111211333213 |
| 313113311333313 | 313122322332213 |
| 112131111233311 | 312112222233212 |
| 331311313331133 | 112121312231112 |
| 313113311333313 | 313113311331313 |
| 322213213222223 | 313113311333313 |
| 123233211232313 | 223223222222222 |
| 313113311333313 | 313112211333313 |
| 313113211333313 | 313111111333313 |
| 312222222232222 | 311111113213113 |
| 111131113331111 | 313112911332393 |
| 323121111333313 | 313113311332313 |
| 312113311333313 | 213122331332212 |
| 313113311333313 | 313113318333313 |
| 113133212213131 | 313113311333313 |
| 211129913231123 | 313113311931313 |
| 319193319333313 | 313119811333313 |
| 313113311331313 | 311113111333313 |
| 393113311333313 | 312219811391318 |
| 333123213333323 | 313113311333313 |
| 313113311331313 | 312113311313333 |
| 112133331233311 | 313113311333313 |
| 131211112223233 | 313113311333313 |
| 233122321333312 | 313113311239323 |
| 313113311333313 | 223228822222222 |
| 313113311333313 | 313333111311313 |
| 313112213233213 | 319119919333313 |
| 313118911333313 | 313113311333393 |
| 131331131211131 | 333113311333313 |
| 311111213321113 | 331311113311113 |
| 313113311322313 | 319113913331313 |
| 313113311332323 | 313113311233313 |
| 313112211333313 | 313113311333313 |
| 313113311233313 | 313113111333313 |

GRUPE 4 : DE ANTWOORDEN OP 15 VRAGEN
(ZIE TABEL 4.1.2.) VAN 80 ZUSTERS

| | |
|-----------------|-----------------|
| 312313211332313 | 323122221333313 |
| 313113311233313 | 232232233231231 |
| 322223221322323 | 322229212322223 |
| 322321113321133 | 313313312321311 |
| 313113321333313 | 222222222233313 |
| 333313211331313 | 313113311333313 |
| 312113212231313 | 312113212223313 |
| 313113313333313 | 312119811333313 |
| 312113211333313 | 323218912232223 |
| 312113212232313 | 311123211333313 |
| 313113311333313 | 322112212232322 |
| 313113311333313 | 312113112333313 |
| 323213313331313 | 312322113331213 |
| 312213213331213 | 313113311332313 |
| 313212211373323 | 319119811311313 |
| 311122222222223 | 313119912333313 |
| 313113311331313 | 313113311333313 |
| 333113311233313 | 313213311233313 |
| 322212222351323 | 231312212221212 |
| 313111113332113 | 313113113311333 |
| 313112213333313 | 312113211331313 |
| 313133311233313 | 312113212221223 |
| 313113311331313 | 319118918333313 |
| 313119898523313 | 238222923222212 |
| 323113212333313 | 311113111311313 |
| 322219811321213 | 323113311232313 |
| 313313313331312 | 311119118393813 |
| 313113311322313 | 313113311333313 |
| 382228993231213 | 222222221232222 |
| 323112212331323 | 313113311311313 |
| 312113311223313 | 313119811331313 |
| 313113311333313 | 313113311333313 |
| 319129821233313 | 222222221232223 |
| 311113221232223 | 323218913321113 |
| 312112212231213 | 313113311333313 |
| 323113211331313 | 313113311333312 |
| 312222211232213 | 333111311331133 |
| 212113311333313 | 222112211232212 |
| 311112212331313 | 313313311231113 |
| 333113311331313 | 111221223212132 |

GRUPEP 5 : DE ANTWOORDEN OP 15 VRAGEN
(ZIE TABEL 4.1.2.) VAN 80 ZUSTERS

| | |
|------------------|-----------------|
| 113113311233313 | 313112213333213 |
| 311131111313313 | 313113311331313 |
| 323113311223313 | 312113311223313 |
| 312119812333213 | 313113311333313 |
| 332112333331113 | 313113211331313 |
| 133331333331312 | 313113311333313 |
| 313113311233313 | 313113311333313 |
| 313112212229898 | 313113311333313 |
| 312112211333213 | 313113311333313 |
| 334322333331313 | 313113212233313 |
| 313113311333313 | 213213212333313 |
| 312123312333313 | 313113311233313 |
| 313112212333313 | 322222222232223 |
| 312313212331313 | 311112113313313 |
| 213122221933313 | 322213211222313 |
| 313113311332313 | 312212211221213 |
| 333332231831123 | 333133313312213 |
| 312113211333313 | 333313211221313 |
| 313113211333313 | 222222222222222 |
| 3131133112331213 | 322112311333213 |
| 323112212233313 | 311113311333213 |
| 313113311333313 | 313113311333313 |
| 322212232311213 | 313113311333313 |
| 313111111331113 | 313112211232323 |
| 322222221332313 | 321323323331313 |
| 313113211333313 | 313113311333313 |
| 223323321231312 | 313119918331313 |
| 132233311222321 | 312122212221212 |
| 113133211229318 | 313113311332213 |
| 232122222331313 | 222229813222222 |
| 131321193331199 | 113113311331113 |
| 312213812322323 | 312212213222222 |
| 111119813333113 | 313113311333313 |
| 312113211333313 | 311113113311133 |
| 313113311333313 | 319113811331313 |
| 313312911333213 | 113123311232212 |
| 223122221333312 | 113113311331113 |
| 313113311333313 | 313118918333313 |
| 322212212322223 | 313113311333313 |
| 313313311332313 | 312113222333213 |

GRUPE 6 : DE ANTWOORDEN OP 15 VRAGEN
(ZIEF TABEL 4.1.2.) VAN 60 ZUSTERS

| | |
|-----------------|-----------------|
| 313113311333313 | 313113311333313 |
| 313129811331313 | 323213211211313 |
| 313113312332313 | 323213212332313 |
| 312211832333313 | 313113311333313 |
| 313113311333313 | 312313221232213 |
| 313119813223313 | 312212212399313 |
| 313113311333313 | 322213311332313 |
| 312113211333313 | 319313311331313 |
| 311132112333213 | 112133311233211 |
| 313113318333313 | 313113311338313 |
| 311113312331313 | 133333311131311 |
| 133131331331211 | 232131213322121 |
| 313313111331313 | 333113312333313 |
| 393113312333313 | 313112311333313 |
| 313113311932313 | 313113311333313 |
| 292219121221213 | 312113213333313 |
| 333312221231312 | 312113211333313 |
| 311228922223222 | 313113311333313 |
| 313112211233313 | 322122212232213 |
| 312118911333313 | 319113311333313 |
| 112392223421222 | 313193311393313 |
| 413212212333313 | 211111113331113 |
| 132339999321131 | 313213311332313 |
| 312113313232232 | 313123311331313 |
| 122112311231212 | 122112212331313 |
| 313113311233313 | 313113311332313 |
| 313113311333313 | 313112211333213 |
| 323313211231313 | 313123311231313 |
| 333313311231313 | 333923322331312 |
| 313113311332313 | 332211212331123 |

GRUPE 7 : DE ANTWOORDEN OP 15 VRAGEN
(ZIE TABEL 4.1.2.) VAN 60 ZUSTERS

| | |
|-----------------|------------------|
| 513115511523313 | 313113511351313 |
| 313133353113313 | 313113311333313 |
| 33331811231132 | 312113311231313 |
| 112222113111111 | 312223321322222 |
| 313213311321323 | 313119999333313 |
| 312113211331313 | 313113311333313 |
| 212133233133311 | 3131133112333313 |
| 212222211223212 | 223213312332323 |
| 213222222233222 | 118239232236911 |
| 313113312333323 | 312113813352313 |
| 313113311333333 | 313113311853313 |
| 212119813333237 | 313132333233111 |
| 313113311331313 | 513113311331313 |
| 313113311233313 | 313113311333313 |
| 313113311333313 | 311113211333313 |
| 312112315232213 | 133331133133111 |
| 313212213332315 | 922112222332382 |
| 322222222222222 | 319113211322313 |
| 312113312331343 | 313113211333313 |
| 222222222232222 | 311313312333313 |
| 313113311233313 | 313313311531313 |
| 333311111331311 | 313313311331313 |
| 313113312233323 | 333212211231313 |
| 333113211331313 | 312122211323213 |
| 323113311231313 | 313113311331313 |
| 313513311231313 | 311113213233333 |
| 313113311333313 | 312112211233313 |
| 321212211331213 | 333113312333333 |
| 222222222223222 | 22222222223212 |
| 323323321311313 | 319228982332223 |

OPGEP 8 : DE ANTWOORDEN OP 15 VRAGEN
(ZIE TABEL 4.1.2.) VAN 40 ZUSTERS

313113311333313
 312112113932223
 131139893333113
 133333133331331
 332311213311113
 312112212232213
 313113311333313
 313113911233313
 313113311933313
 322112211332313
 322222222322323
 313113311333313
 313113311233313
 313923311231313
 313113311333313
 133331311331312
 323222311233323
 313113311239113
 313113311333313
 313219916331313
 313112311333313
 312223322233322
 113113311231313
 313113311333313
 312112918333393
 312122211233213
 319112212233213
 311113311233313
 313113311331123
 313133211312313
 282222222839383
 319113311331313
 312121212223212
 313119811333313
 132332139211133
 232331923221122
 322222222333223
 312321331331213
 323219822232322
 313113311313313

GRUP 9 : DE ANTWOORDEN OP 15 VRAGEN
(ZIE TABEL 4.1.2.) VAN 40 ZUSTERS

31321332122312
 31311222222222
 312112988331213
 31311331123313
 31321331333323
 313113211433313
 233113311331313
 233113311331313
 123388311131319
 313118911332313
 222212212331213
 313119811313313
 313113211321313
 322212223221211
 313113311331313
 212123322231212
 312112212333413
 313113311333313
 313113311333313
 313113311322313
 313119911933313
 313183911883323
 313112311333313
 313119811333113
 222222222222222
 313118911332323
 312123212223313
 319212313323213
 313113311233313
 222222222223313
 2222222222231312
 212112221231222
 313113311333213
 313113311333313
 313311113331313
 322122222331113
 333313111331313
 313113311333313
 331111113331313
 313113311233313

GROEP 10 : DE ANTWOORDEN OP 15 VRAGEN
(ZIE TABEL 4.1.2.) VAN 40 ZUSTERS

312112211333313
 232221223322212
 32322222231393
 313313313311313
 313213321231313
 313112212333313
 323113312322323
 22232222331222
 313113311333313
 212112231233213
 313113311333313
 312112212333313
 311113311333313
 312219311231313
 331313211312333
 313113311333313
 313113311333313
 333311211331113
 313113311333313
 322213212333213
 311113113333313
 312113311333313
 212122211233212
 338313911331813
 333313212331313
 393913311231313
 312112212233213
 312112212232223
 323229822332212
 13313111331111
 312112212333313
 313313313221222
 313113312333313
 313113311333313
 22322222232212
 313113311333313
 913112211223323
 313113311331383
 313113211333313
 312113311231313

BIJLAGE II.3. DATA BEROEPEN(ZIE PARAGRAAF 5.2.1.)
 18 BEROEPSEISEN VAN 47 BEROEPEN
 CODERING 1: BEROEPSEIS RELEVANT
 CODERING 0: BEROEPSEIS NIET RELEVANT

```

1 001110000000100000
2 001100000000100000
3 000000101000001010
4 000100000010010000
5 000001010010000001
6 110000110000000000
7 110000001100000000
8 000100001100100000
9 100100001000100000
10 011000001000010000
11 001100000000101000
12 001100000100100000
13 000001110000000001
14 000010000000011010
15 001000000010100010
16 000000000010011100
17 001010000000000010
18 000001010001000001
19 000010000010010100
20 000001110000000001
21 001010000000011010
22 000110000010010000
23 000110000000010000
24 001000000101001000
25 000001010000000101
26 100000001100001000
27 101001001000000000
28 101000000100000000
29 000000110001000001
30 010000000001001100
31 010000101100000000
32 000000101101000000
33 000000000001010110
34 000000010100100010
35 000001110001000000
36 110000000001001000
37 001100000100100000
38 010000001100001000
39 000001000000100011
40 000000100000011110
41 010000100001001000
42 100100001100000000
43 110000000100100000
44 000010000010010100
45 100100000000100001
46 000010000010010100
47 001100000010010000

```


BIJLAGE II.4. DATA OPSTELLEN(ZIE PARAGRAAF 5.2.2.)
57 KENMERKEN VAN 181 OPSTELLEN
CODERING 1 : KENMERK AANWEZIG
CODERING 0 : KENMERK AFWEZIG

VFRVOIG

[illegible]

BIJLAGE II.4. DATA OPSTELLEN(ZIE PARAGRAAF 5.2.2.)
67 KENMERKEN VAN 181 OPSTELLEN
CODERING 1: KENMERK AANWEZIG
CODERING 0: KENMERK AFWEZIG

VERVOLG

[illegible]

BIJLAGE 11.4. DATA OPSTELLEN (ZIE PARAGRAAF 5.2.2.)
67 KENMERKEN VAN 18) OPSTELLEN
CODERING 1: KENMERK AANWEZIG
CODERING 0: KENMERK AFWEZIG

VERVOLG

[illegible]

BIJLAGE II.5. DATA A.M.-ONDERZOEK
(ZIE PARAGRAAF 4.5.)
DE NUL EN EEN CODERING VAN DIT ONDERZOEK
STAAT IN TABEL 4.5.2.

Bijlage III. Populatie factorstructuur simulatie-onderzoeken

Simulatie I

Populatie factorstructuur: geen factoren

Simulatie II

Populatie factorstructuur: geen factoren

Simulatie III

Populatie factorstructuur: één algemene factor met gelijke ladingen van 0,50 voor 30 operatoren

Simulatie IV

Populatie factorstructuur (nulwaarden zijn weggelaten):

| | | | | | |
|----|------|-------|------|-------|-------|
| 1 | 0,40 | | 0,60 | | |
| 2 | 0,40 | | | 0,55 | |
| 3 | 0,40 | | | | 0,60 |
| 4 | 0,45 | 0,55 | | | |
| 5 | 0,45 | | | -0,60 | |
| 6 | 0,45 | | | | 0,50 |
| 7 | 0,50 | | | 0,60 | |
| 8 | 0,50 | -0,60 | | | |
| 9 | 0,50 | | 0,50 | | |
| 10 | 0,55 | | | 0,50 | |
| 11 | 0,55 | | | | 0,50 |
| 12 | 0,55 | | | | |
| 13 | 0,60 | | | | |
| 14 | 0,60 | 0,40 | | | |
| 15 | 0,60 | | | 0,45 | |
| 16 | 0,65 | | | | |
| 17 | 0,65 | | 0,50 | | |
| 18 | 0,65 | | | | |
| 19 | 0,70 | 0,40 | | | |
| 20 | 0,70 | | | 0,45 | |
| 21 | 0,70 | | | | -0,45 |
| 22 | 0,75 | | | | |
| 23 | 0,75 | | | -0,40 | |
| 24 | 0,75 | | | | |
| 25 | 0,80 | | | | |
| 26 | 0,80 | | 0,40 | | |
| 27 | 0,80 | | | | |

Simulatie V

Populatie factorstructuur: twee gelijke groeps-
factoren (zie tabel 6.2.2.)

Simulatie VI

Populatie factorstructuur (nulwaarden zijn weggelaten)

| | | | |
|----|------|-------|-------|
| 1 | 0,60 | | |
| 2 | 0,50 | -0,60 | |
| 3 | 0,40 | | -0,50 |
| 4 | 0,50 | | 0,50 |
| 5 | 0,60 | | 0,50 |
| 6 | 0,80 | | |
| 7 | | 0,60 | 0,50 |
| 8 | | 0,50 | -0,60 |
| 9 | | 0,40 | 0,70 |
| 10 | | 0,50 | 0,70 |
| 11 | | 0,60 | |
| 12 | | 0,80 | |
| 13 | | | 0,80 |
| 14 | | | 0,70 |

Simulatie VII

Populatie factorstructuur: zie Simulatie VI

Simulatie VIII

Populatie factorstructuur: drie groepsfactoren en
drie subfactoren (zie tabel 6.3.2.1.).

stellingen

1. De iteratieve clusteranalyse is een bruikbare methode voor schaalconstructie en type-onderzoek.
2. Het informatieverlies, dat door nul en één coderingen optreedt, is psychologisch meestal niet relevant.
3. Bij het afbeelden van psychologische 'overeenkomsten' in 'afstanden' wordt vaak onvoldoende rekening gehouden met de wenselijkheid alle eigenschappen van 'afstanden' in termen van 'overeenkomsten' te interpreteren.

C.H. Coombs
A theory of data
New York: Wiley, 1964

4. Chefs kunnen geen goede promotievoordrachten van hun directe medewerkers maken.
5. Het dient bij bedrijfspsychologische selectie de gewoonte te worden om sollicitanten in de gelegenheid te stellen hun sollicitatie in te trekken na kennisneming van het psychologisch rapport.

A.C. Rümke
Nieuwe selectie-procedures bij de P.T.T.
De Psycholoog, 1968, 3, 150-152

6. Diegenen, die selectie een noodzakelijk kwaad noemen, ontkennen de positieve functie van de menselijke rivaliteit.

W. Hofstee
Selectie: een noodzakelijk kwaad?
Congresboek van Groningse Psychologen Congres

7. De reden dat europese meisjes in tegenstelling tot europese jongens geen beroemdheden van het lichte genre als ego-ideaal beschrijven, moet meer worden toegeschreven aan de narcistische houding, die in onze cultuur van de vrouw wordt vereist, dan aan het sterker normbesef, dat haar wordt bijgebracht.

G. Kempen en A.H. Boon van Ostade

Een typologie van ideaalbeelden bij europese jongeren

Ned. T. Psychol., 1969, 24, 46-60

8. De ontwikkeling van de bureaucratische staatsvorm in Rusland heeft tot gevolg, dat alleen een tweede marxistische revolutie nog de communistische heilstaat kan brengen.

9. Aan de universiteit moeten de studenten het laatste woord hebben ten aanzien van het onderwijs, het administratief- en technisch personeel ten aanzien van de efficiency en de academische staf ten aanzien van het wetenschappelijk onderzoek.

10. De betekenis van het woord 'subjectief' voor menselijke waarnemingen is het beste te leren door zich te abonneren op twee verschillende nederlandse dagbladen.

11. De factoranalytische methode kan als een Rorschach Test worden gebruikt.

12. Psychologen dienen lid te zijn van het N.I.P.

Nijmegen, 23 november 1969.

A.H. Boon van Ostade.

stellingen

1. De iteratieve clusteranalyse is een bruikbare methode voor schaalconstructie en type-onderzoek.
2. Het informatieverlies, dat door nul en één coderingen optreedt, is psychologisch meestal niet relevant.
3. Bij het afbeelden van psychologische 'overeenkomsten' in 'afstanden' wordt vaak onvoldoende rekening gehouden met de wenselijkheid alle eigenschappen van 'afstanden' in termen van 'overeenkomsten' te interpreteren.

C.H. Coombs
A theory of data
New York: Wiley, 1964

4. Chefs kunnen geen goede promotievoordrachten van hun directe medewerkers maken.
5. Het dient bij bedrijfspsychologische selectie de gewoonte te worden om sollicitanten in de gelegenheid te stellen hun sollicitatie in te trekken na kennisneming van het psychologisch rapport.

A.C. Rümke
Nieuwe selectie-procedures bij de P.T.T.
De Psycholoog, 1968, 3, 150-152

6. Diegenen, die selectie een noodzakelijk kwaad noemen, ontkennen de positieve functie van de menselijke rivaliteit.

W. Hofstee
Selectie: een noodzakelijk kwaad?
Congresboek van Groningse Psychologen Congres

7. De reden dat europese meisjes in tegenstelling tot europese jongens geen beroemdheden van het lichte genre als ego-ideaal beschrijven, moet meer worden toegeschreven aan de narcistische houding, die in onze cultuur van de vrouw wordt vereist, dan aan het sterker normbesef, dat haar wordt bijgebracht.

G. Kempen en A.H. Boon van Ostade

Een typologie van ideaalbeelden bij europese jongeren

Ned. T. Psychol., 1969, 24, 46-60

8. De ontwikkeling van de bureaucratische staatsvorm in Rusland heeft tot gevolg, dat alleen een tweede marxistische revolutie nog de communistische heilstaat kan brengen.

9. Aan de universiteit moeten de studenten het laatste woord hebben ten aanzien van het onderwijs, het administratief- en technisch personeel ten aanzien van de efficiency en de academische staf ten aanzien van het wetenschappelijk onderzoek.

10. De betekenis van het woord 'subjectief' voor menselijke waarnemingen is het beste te leren door zich te abonneren op twee verschillende nederlandse dagbladen.

11. De factoranalytische methode kan als een Rorschach Test worden gebruikt.

12. Psychologen dienen lid te zijn van het N.I.P.

Nijmegen, 23 november 1969.

A.H. Boon van Ostade.

